EXISTING INFORMATION

Before planning the information collection activities of step one, we evaluated the geographic information already available for the region [7]. The entire region had been covered by black-and-white aerial photographs—of various dates between 1957 and 1974 and at scales between 1:30 000 and 1:70 000—which had been used to produce 1:50 000 scale topographic maps. There were also various small-scale base maps, thematic maps of climate, geology, vegetation and soils, and geo-referenced tabular data.

A land and water resource survey report for Arusha Region was also available [20]. The report includes Landsat imagery of the entire region with interpretation overlays showing areas with potential groundwater (based primarily on geologic lineations) and dividing the region into land units (based on vegetation and topography). In the report, related land units are grouped into 36 land systems. Most of the land systems include a great diversity of vegetation, soils and slope angles, making them too general for specific planning. On the other hand, the component land units are too specific in most cases to serve as planning units.

We decided that the land unit boundaries of the survey report needed to be verified against new data and then regrouped into larger units suitable for the regional planning process. Given the size of the region and difficulty of overland travel, combined with our time and budget constraints, the regional planners agreed to a multi-level rapid reconnaissance approach which depended primarily on enhanced Landsat imagery and secondarily on low-level aerial survey and analysis of topographic maps and existing reports.

STEP ONE: RAPID RECONNAISSANCE

Four separate data sources were used in step one. The first source was Landsat imagery enhanced by digital processing to improve visual interpretation. The second was a set of general slope angle maps derived from analysis of the existing 1:50 000 scale topographic maps. The third was a low-level aerial survey, or systematic reconnaissance flight (SRF), in which a pilot and three observers flew at 91 m (300 feet) above ground level in a high-wing aircraft. The observers counted point data on the ground (livestock, wildlife, huts, etc) and estimated the areal percentages of various types of natural vegetation and land use. The fourth data source was the set of existing reports and maps from ground surveys of sub-regional areas and resource variables.

LANDSAT IMAGERY

The primary source of information for step one was Landsat imagery. Of the nine Landsat scenes covering Arusha Region, five included more than 90 percent of the area. The Environmental Research Institute of Michigan (ERIM) was contracted to digitally enhance these five scenes, indicated by date and scene identification number below:

25 December 1972 (1153-07131 and 1155-07133)
4 September 1973 (1408-07174)
24 January 1976 (2367-06582)
25 January 1976 (2368-07040).

Enhanced false colour composite (FCC) photographic products were generated for maximum clarity, resolution and colour separation to improve information on land cover [12]. The primary enhancement was for contrast accentuation, especially those digital values related to land cover. The final processing steps involved careful colour balancing and registration. An example of the results is shown in Figure 1. Standard FCCs were used for the small areas not covered by the five enhanced scenes.

Because of Landsat's coarse resolution, topography cannot be judged accurately from the imagery alone. Moreover, because the colours and patterns of the imagery are difficult to associate with ground features without actually checking data from selected areas on the ground against the same areas represented on the imagery, the Landsat data must be supplemented by other information about topography, soil and vegetative cover, and such cultural features as roads, villages and cropland.

SLOPE ANGLE ANALYSIS

Topographic information was collected to aid interpretation of the Landsat imagery. Slope angle information, in particular, was important for land capability assessment [8, 9, 23]. The information was acquired from the 15 minute topographic maps published by the Tanzanian Division of Surveys and Mapping (Series Y-742). Arusha Region is covered by 137 maps at a scale of 1:50 000 [4]. The contour lines of these maps were used to divide the region into general slope angle mapping units of five classes: 0 to 3 percent, 3 to 10 percent, 10 to 25 percent, 25 to 50 percent, and 50 or greater percent slope.

Slope angle was determined rapidly by taking the ratio of vertical distance to horizontal distance, as measured from the spacing of contour lines on the topographic maps [13]. Delineations of homogeneous slope angle mapping units were then made on vellum overlays. Since the slope angle information was to be integrated with the Landsat imagery at 1:250 000 and 1:500 000 scales, initial delineations of mapping units were generalized and quite large (often greater than 100 cm^2) on the 1:50 000 scale maps to allow the required reductions in scale.

Occasionally, a mapping unit consisted of a complex of two slope angle classes. In these instances, the

FIGURE 1 (opposite page) *Bottom:* Enhanced Landsat false colour composite of northern Tanzania and southern Kenya taken in January 1976 (Mt Kilimanjaro on the right, Mt Meru and Arusha Region, Tanzania in the lower left). *Center:* Mosaic of natural colour vertical aerial photographs of the outskirts of Arusha Town in Arumeru District of Arusha Region, taken February 1980. *Top:* Natural colour oblique aerial photo of Arumeru District, taken January 1980 (Landsat image enhanced by Environmental Research Institute of Michigan, photography by Dana M Slaymaker, design by Paul Mirocha).

Pictometry EX2002 (Part 3 of 13)
Roofr v. Pictometry
IPR2023-00436
0769

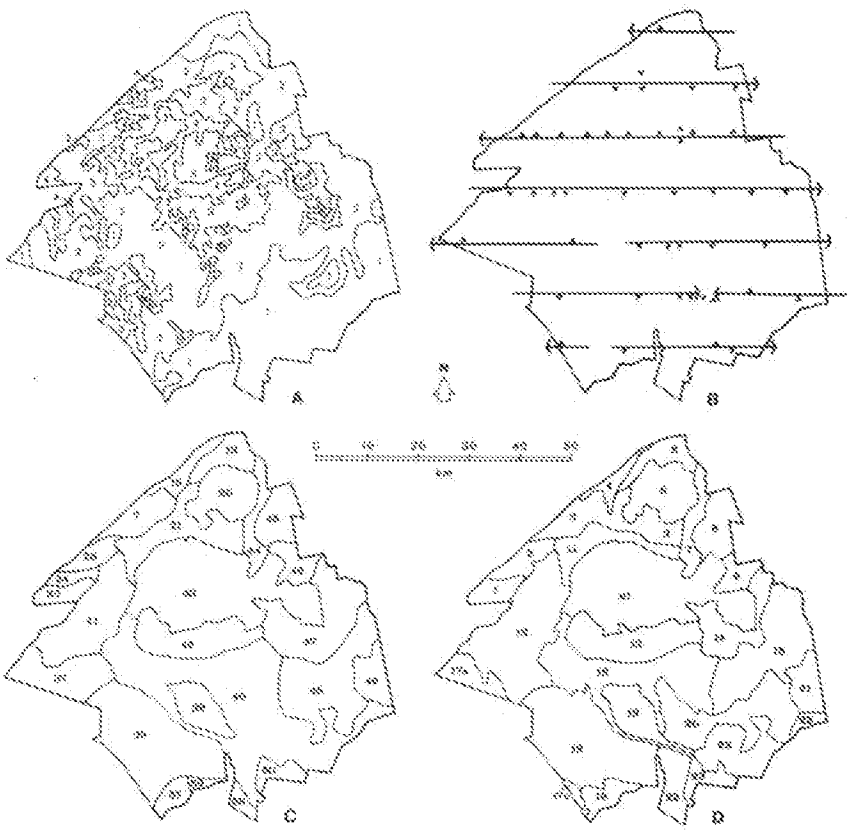


FIGURE 2 Arumeru District, Arusha Region, Tanzania

FIGURE 2A Slope angle mapping units
Class 1 = 0–3 percent, class 2 = 3–10 percent, class 3 = 10–15 percent, class 4 = 25–50 percent, class 5 = greater than 50 percent slope

FIGURE 2B SRF flight lines and photograph locations
Long arrows show direction of flight east or west, triangles show location and direction of oblique 35 mm photographs

FIGURE 2C Preliminary land units
The unit numbers are part of a regional numbering system

FIGURE 2D Revised land units
The unit numbers are part of a district numbering system

delination was identified by the two respective class categories. For example, 1/4 indicated a complex of class 1 (0 to 3 percent slope—relatively level) and class 4 (25 to 50 percent slope—relatively steep). Such a complex might indicate a steeply dissected planar surface. The individual channel banks would be far too narrow for separate delineation. Omitting the channels would leave the impression of a smooth level surface (eg, a dry lake bed). By describing the unit as a complex, the analyst was subsequently able to take into consideration the frequent steep-sided channels.

The 1:50 000 slope angle maps were reduced to 1:250 000 and 1:500 000 scales and were composited for ease in working with the other data (Figure 2A).

LOW-LEVEL AERIAL SURVEY

Interpretation of the colours and patterns seen on the Landsat imagery was aided by data from low-level aerial surveys using the SRF technique. The SRF was originally designed for studies of population size, density and distribution, and movements of large wildlife and domestic livestock [16]. The method has recently been modified and developed for collecting both point data (numbers of wildlife, livestock, houses, etc) and cover data (percentage of ground covered by various types of topography, soil and natural vegetation, and land use, including various crop covers) [15, 22]. With SRF, a framework of grid squares is superimposed on a map of the region and then each grid square is sampled. The flight lines were spaced across the region so that the aircraft passed along the middle of each row of squares. Data were collected systematically by low-level aerial observation and photography within each square along each flight line.

In a period of three weeks, a total of 85 500 km^2 —including all of Arusha Region—were surveyed by two aircraft (a Cessna 182 and a Cessna 206) using standard SRF methods [10]. The four person crews in the two aircraft were drawn from EcoSystems Ltd, the Tanzania Ministry of Natural Resources and Tourism, and the University of Dar es Salaam. The flight lines were oriented east-west to transect the major geographic features of the region and were spaced at 10 km intervals. The aircraft were flown at approximately 91 m (300 feet) above ground level at a mean ground speed of 168 km per hour. The sampling strips averaged 304 m in width to give an overall sample of 3.04 percent. All observations were recorded in successive, one minute time sample units along each flight line. Each sample unit was approximately 2.8 km in length at the mean ground speed of the aircraft.

Radar altimeters helped control the height of both aircraft. In the Cessna 206, navigation and map location of the sample units were carried out with the aid of a global navigation system (GNS 500A) that tunes into the worldwide network of VLF/Omega radio beacons. This aircraft was used where landmark aids to navigation were relatively sparse. In the Cessna 182, navigation was more conventional with the aid of maps and visual reference to ground features. Each sample unit was assigned to a

0771

10 x 10 km grid square so that the sample unit data could be analyzed by computer either individually or aggregated to characterize whole grid squares. The data were presented as grid maps (Figures 3 and 4) and as tables showing aggregate data by district (Tables 1 and 2).

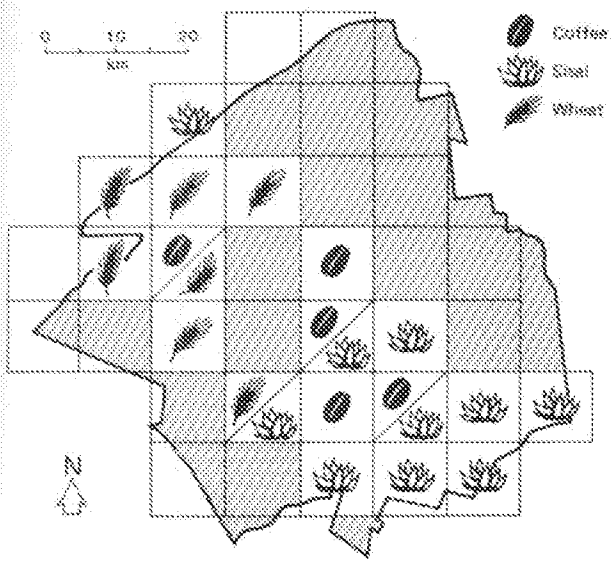


FIGURE 3 SRF crop data for Arumeru District
The map indicates in which 10 x 10 km grid squares the airborne observers saw at least some of each cash crop.

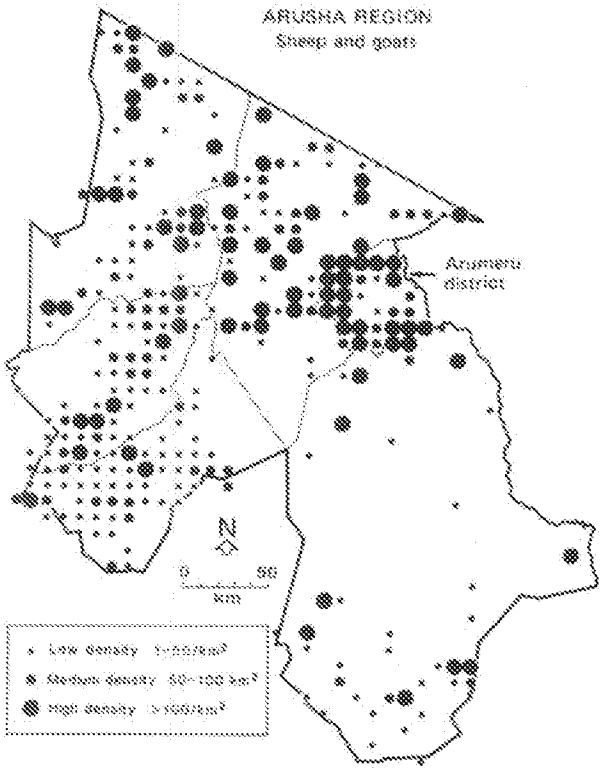


FIGURE 4 SRF livestock data for Arusha Region
The circles correspond to those 10 x 10 km grid squares in which at least some sheep or goats were seen by the airborne observers.

TABLE 1 SRF crop data for Arusha Region and its six districts

Crop	NG	MB	MN	MD	KT	AR	RG
Maize	4	55	61	18	6	100	32
Banana	3	30	38	3	1	73	8
Marrow (pumpkin)	1	19	29	3	3	14	8
Pigeon peas	1	12	13	12	1	38	7
Castor oil	1	10	16	3	2	14	5
Tobacco		23	3	1		3	4
Eucalyptus	1	14	6	1		33	4
Beans	1	12	5	2		16	3
Cassava	1	10	2	2		17	3
Wheat	1	7	3	5		21	3
Mango		8	6	1		33	3
Sugarcane		6	2				2
Coffee		5	2			17	1
Sisal			1	1		28	1

The numbers show the percentage of sample units in which the airborne observers saw at least some of each crop (each unit is one minute of flight or approximately 2.8 km). NG = Ngorongoro District; MB = Mbulu District; MN = Hanang District; MD = Monduli District; KT = Kiteto District; AR = Arumeru District; RG = Arusha Region.

TABLE 2 SRF cattle data for Arusha Region and its six districts

Area	Density per km²	Number of animals	Percentage of regional total
Ngorongoro District	16.35	243 280	11
Mbulu District	42.92	337 485	15
Hanang District	49.27	469 068	21
Monduli District	22.73	363 253	16
Kiteto District	19.83	650 943	29
Arumeru District	80.56	212 232	8
Arusha Region	37.18	2 276 184	100

The least significant difference (LSD) between any two districts' mean cattle densities is 8.11. The 95 percent confidence interval for the regional estimate of total number of cattle is ± 6 percent (19).

In addition to visual observations of point and cover data, one observer in each aircraft took an oblique 35 mm photograph (colour slide on Ektachrome ASA 400 film exposed through a 50 mm lens) of the landscape once every three minutes. The photographs were located along the flight lines (Figure 2B) for later registration on the Landsat imagery and base map. Approximately one oblique landscape photograph was taken for each 10 x 10 km grid square. One of the oblique photos is included in Figure 1.

GROUND SURVEYS

Although extensive ground data collection was impractical, we did some ground survey work to acquaint ourselves with the landscapes of Arusha Region. A soil scientist and two range scientists also carried out brief regional surveys [3, 17] that provided additional current information. During a three month period, the regional government completed a questionnaire survey of approximately one-third of Arusha Region's villages. Considerable information on land resources and land use (particularly range resources, livestock management and crop production) and social, economic, and political informa-

0772

tion about each profile village was collected [18]. Unfortunately, the volume and variety of data from the village profiles impeded manual analysis. A minicomputer was eventually installed in the regional offices to handle the data management and analysis problem, but the village profile data were not available in time to be integrated with the rapid reconnaissance of step one.

INTERPRETATION AND USE

The rapid reconnaissance required an integrative approach to land resource analysis. The primary data source was the Landsat imagery. Additional information included the slope angle maps, 35 mm oblique aerial photographs and point and cover data from the low-level aerial survey, ground observations, and existing reports and maps. The Landsat imagery and supplemental information were integrated relatively simply by registering both to a common base map at 1:500 000 scale (the regional map of villages). We began by delineating on the imagery distinct areas of relatively homogeneous colour, texture and pattern, and other recognizable image characteristics. We then tested these delineations against the other information to determine whether the imagery units should be combined or further subdivided. The result was a regional set of approximately 250 preliminary land units that appeared to be both internally homogeneous and distinct from neighboring units in terms of topography, vegetation and land use (Figure 2C).

These land units were further examined, modified where necessary, and verified by officials with local knowledge of the land being mapped. Successful application of remote sensing technology to resource inventories depends on the ability of image interpreters to associate image characteristics with the land resources and uses being mapped [13]. Rather than depending solely on our own abilities and the supplemental information that provided most of our knowledge of the region, we involved district level land management officials (primarily agriculture, livestock and forestry officials) in two to three day district workshops. The first half-day was spent familiarizing the workshop participants with the Landsat imagery covering their district and the appropriate ancillary information. We then superimposed on the imagery our preliminary land units with a clear acetate overlay. With their working knowledge of the land and people of the areas in question, the participants repeated the process of testing the land units against the imagery and other information. They frequently changed the alignment of land unit boundaries and combined or subdivided land units to produce a set of revised land units for each district (Figure 2D).

Land units were relatively stable physical entities and the basic building blocks for land planning. They could be grouped in various ways to suit particular plannings needs: the agricultural planner would group them into planning units quite different from the groupings chosen by the livestock planner. For general planning purposes, however, the workshop participants built up land planning units by grouping those land units that included the same or related land uses and ethnic groups. The grouped land units had common resources, problems and potentials which were distinct from those of neighboring land planning units. A great deal of subjective judgment was involved; hence, active participation by the intended users of the planning units was needed. The final results were six district sets of land planning units (Figure 5) averaging approximately 11 planning units per district, which was a manageable number for district level planning.

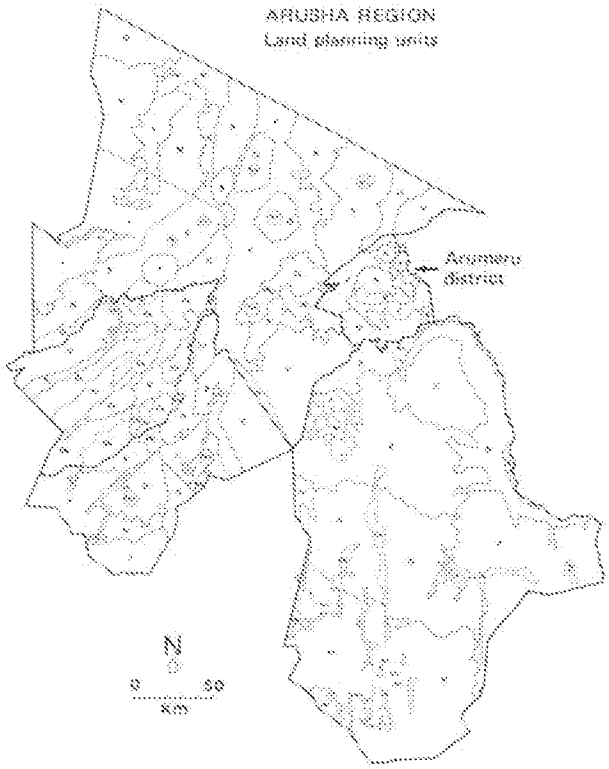


FIGURE 5 Land planning units of Arusha Region
Each of the six districts has its own set of units with a distinctive lettering system, with the exception of Monduli and Ngorongoro where the boundary between these recently separated districts had yet to be determined (northern part of the region).

We summarized the point and cover data collected by the low-level aerial survey for each of the land planning units in order to assist district officials in identifying needs or potentials for development. Table 3 shows examples of data from two planning units that appear superficially similar but that, according to the survey data, are different in important socio-economic features such as degree of agricultural mechanization, percentage of houses with metal roofs (rather than thatch) and the ratio of cattle to sheep and goats.

The district planning teams made different uses of the land planning units. Each team was responsible for selecting geographic areas in which to concentrate development activities and for developing the rationale for their selection. One team designated a single land planning unit as its area of concentration. Four other teams

0773

decided to concentrate on villages representative of several land planning units covering a diverse range of agro-ecological conditions within each district. One team decided to concentrate on the most accessible land planning units. In each of the six districts, a subset of the district's land planning units was given priority for development. The selections were endorsed by the political and administrative leaders of the district and the region [1].

TABLE 3 SRF data summaries for two district land planning units

	Arumeru District Land Planning Unit J	Hanang District Land Planning Unit E
AREA (km^2)	622	468
Number of sample units	22	16
VEGETATION (percent)		
Agriculture	50	31
Grassland	50	13
Bushed grassland		25
Wooded bushland		13
SOILS (percent)		
Brown	77	
Red-brown		50
Grey		25
Grey-brown		25
LIVESTOCK		
Density per km^2	371	153
Ratio (cattle: sheep/goats)	0.5:1	2.5:1
WILDLIFE		
Density per km^2	1.1	2
HOUSES		
Density per km^2	21.8	13
Metal roof (percent)	38	15
CROPS		
Cultivation methods		
Mechanical	60	27
Ox-drawn	40	
Hand		73
Dominants	Primarily maize; Secondarily wheat, bananas, beans, and coffee	Primarily maize; Secondarily marrows and tobacco
Diversity	14 species (including marrows, pigeon peas, castor beans, eucalyptus, cassava, mango, wattle, grevillea, potatoes)	12 species (including bananas, pigeon peas, castor beans, eucalyptus, wheat, mango, sugarcane, wattle, pine)

Cover data are shown as the percentage of sample units (minutes of flight) in the land planning unit that were dominated by the particular data category (for each minute of flight—covering approximately 2.8 km—the observer crew determined which data category was dominant). Point data are shown as densities. The percentage of all houses that have "modern" metal roofs is also shown. Dominant crops are those most frequently seen from the air. Diversity of crops is given as the total number of crop species seen in the land planning unit. Because the SRF data were originally stored by individual sample units accurately referenced to a base map, it was possible to reorganize the data by computer after land planning units were delineated to show what was found in each unit by the aerial survey.

STEP TWO: DETAILED SURVEY

After a planning unit was given high priority for land development or rehabilitation, more detailed and accurate (and more expensive) survey techniques were justified for project design and implementation. Detailed survey work began with complete vertical aerial photographic coverage of high priority areas and extraction of information from the photographs with the help of people living in the survey areas.

VERTICAL AERIAL PHOTOGRAPHY

Aerial photographs are typically acquired with specialized aerial mapping cameras using 23 x 23 cm (9 x 9 inch) format film. These large-format cameras are manufactured to render metrically correct images from which measurements can be taken with great accuracy, approaching 0.1 m of the correct ground position when measured from photographs taken 1000 m above the ground. Mapping camera systems are very expensive, however, costing more than many of the small aircraft used to carry them. Mapping of land use and other environmental features for planning purposes requires accuracy in the tens of meters rather than fractions of a meter. Small format 70 mm or 35 mm cameras are fully adequate and far less expensive to purchase and use [5]. We chose the 70 mm rather than the 35 mm format because each 70 mm frame covers a larger area with greater resolution.

The 70 mm camera (Hasselblad 500 ELM fitted with a Hasselblad 50 mm Distagon lens) was mounted in a locally hired, single engine Cessna 206 aircraft flown by a local charter pilot. The aircraft was not specially modified for aerial photography (*ie*, no belly port to allow the camera a direct vertical view through the fuselage). The camera was mounted on a Fairchild 283 camera mount attached to rods clamped on the rear seat support rails (seats removed) so that the camera extended through the open cargo door frame (door removed). This arrangement allowed a vertical view from the side of the aircraft. The camera and mount disrupted the normal airstream of the aircraft, so both had to be fitted with air baffles that smoothed the airflow. Even with the baffles, the protruding equipment caused asymmetry in the performance of the aircraft, but this was easily corrected by the pilot. Even these minor modifications to the aircraft, however, required examination and approval by the Tanzania Directorate of Civil Aviation.

The photographer navigated with the aid of a television monitor that was boresighted with the camera. The shutter was released electronically so that each photo frame overlapped the preceding frame by approximately 60 percent (for subsequent stereoscopic viewing). The flight lines were spaced to achieve approximately 30 percent overlap between adjacent lines of photos.

The aircraft was flown at approximately 2576 m (8500 feet) above ground level to give a nominal contact scale of 1:50 000 using Ektachrome MS film 5256 processed to a negative. An automatic printer was used to obtain 5 x 5 inch (± 13 x 13 cm) natural colour enlargements at a nominal print scale of 1:25 000. Natural

colour was used, in preference to black-and-white or colour infrared photography, because natural colour photographs show the most useful detail for relatively dry, sparsely vegetated landscapes.

Twelve wards (village clusters) were selected for complete medium-scale photo coverage. The wards were scattered over the region and varied between 250 and 2500 km^2 in size. Only eight of the wards were completely photographed during a two week period because of problems with weather and aircraft operations. Two prints were provided from each exposure, one as part of a set of separate prints for stereopair viewing, the other as part of a mosaic of prints covering each ward. Mosaics were constructed by matching and overlapping prints to create a continuous picture of the ground, then slicing through the overlap areas with a razor, taping matching ends together on the back, and laminating the whole mosaic to form a single piece. An example of these photomosaics is shown in Figure 1.

VILLAGE DIALOGUE

Just as the participation of district land management officers was essential to the delineation of district land planning units, land planning in villages depended on the active involvement of the people living there. Villagers know a great deal more about their village area than outsiders could ever learn in a reasonable period of time at reasonable cost, and they can be effective partners in the planning process. District officials usually have broad knowledge of specific topics such as crops, livestock or forestry, while villagers usually have specific knowledge of a broad topic--making a living from their village area. Each group has much to offer the other in the process of rural development.

The project experimented with the use of photomosaics of village areas as catalysts for communication between officials and villagers. With a magnifying glass, villagers could identify their own houses and fields, from which they could quickly relate to the photomosaic as an overview of their local area. Once oriented, villagers were able to help district officials locate the boundaries of village areas and identify habitations, land use, natural vegetation types and so on. The mosaics were laminated so that village boundaries and other delineations could be be marked directly on the photographs and later erased and changed. Village development officials found that, in addition to helping organize the collection of geographic information from villagers, the photomosaics gave substance and authenticity to a potentially intangible discussion of land use issues. Much time was saved by assuring that each person was clearly aware of the specific area being discussed. Furthermore, villagers enjoyed working with the photomosaics, which created an atmosphere of acceptance and enthusiasm for the planning process [7].

CONCLUSIONS

We can draw some qualitative conclusions about the efficiency of our approach, even though the project was not designed to monitor the use of the land information materials produced.

The costs of collecting and presenting the land resource information were reasonable--approximately US$0.90 per km^2 for the rapid reconnaissance survey and approximately US$4.00 per km^2 for the vertical colour aerial photographs. We consider our methods highly efficient in terms of area covered, time required, and quantity and quality of data produced for each dollar spent. It is difficult, however, to measure efficiency in terms of total use compared to total cost, including time, effort and money spent in both acquiring and using the information. Measured more simply in terms of meeting the objectives of steps one and two, the efficiency of our remote sensing approach could have been improved.

In step one, a manageable number of land planning units was delineated. These units constituted an important basis for implementing the strategy of geographic concentration in the regional development planning process[1]. The slope angle maps and the oblique landscape slides were invaluable for interpretation of the Landsat imagery. The point and cover data from the low-level aerial survey, however, were not used as expected in the delineation of land planning units or in the assignment of priority to these units. The intended data users seemed to be overwhelmed by the volume of data.

In step two, the colour aerial photographs provided useful details about high-priority areas. The photomosaics facilitated dialogue between officials and villagers. The mosaics and separate prints were also used successfully in the study of existing roads, crop density and coverage, population (settlement) density, delineation of village boundaries, enumeration of informal entrepreneurs in Arusha Town, and sitings of gravity flow water supplies, boreholes and a hydropower facility [7]. The photographs were not used to select representative villages, however, nor were they essential for initiating dialogue with villagers.

Much of the lack-of-use problem can be attributed to insufficient participation by regional staff (government and contract employees) in the detailed planning of the project and, more important, in the collection and interpretation of the information [11]. Time was a critical factor. Nearly all data collection activities occurred within a period of three months--January through March 1980. Regional staff had to approve the survey plans as a single package and then had to assimilate the results as a single package.

The sequence of data collection and interpretation activities could have been reordered and extended to allow greater participation by planning and land management officials. In the future, we recommend the following sequence:

(1) select suitable Landsat imagery, with digital enhancement of scenes for which the standard FCC product shows poor colour balance, poor resolution, striping, etc

(2) analyze slope angle from existing topographic maps

(3) interpret imagery and delineate land units and land planning units with the land management officials familiar with the area

0775

(4) assign tentative priorities to land planning units with the district planning teams, using brief summaries of the land and people in each unit (prepared with the land management officials, based on their knowledge and existing reports)

(5) perform low-level aerial surveys of the priority planning units at 5 to 10 percent sampling intensity, using visual counting of point data and vertical sample photography (35 mm colour) for cover data [23]

(6) select the highest priority planning units with the district planning teams, using the summarized low-level aerial survey data

(7) collect minimum village profile data from all villages in the highest priority planning units

(8) select representative villages with the district planning teams, using summarized village profile data

(9) initiate official-villager dialogue within representative villages

(10) acquire comprehensive colour aerial photographs of representative villages and surrounding areas and

(11) interpret aerial photographs with villagers to meet specific project design and implementation needs

This sequence builds upon our experience in Arusha Region. It should be applicable, with appropriate modifications, to almost any project which requires selection and description of target areas of concentrated activity. Selection of target areas is often dictated by political considerations or by known characteristics of the environment. Political choices, however, may have to be legitimized by application of certain decision criteria to an objective data base, as in Arusha Region. Also, the environment may not be well known, as is often the case in Africa. Existing maps are usually quite inaccurate, and the local knowledge of officials is usually limited to the areas immediately surrounding roads and accessible village settlements. In these cases, it is advisable to generate an up-to-date, comprehensive, reconnaissance level map of land and human resources to assist selection of target areas. If the whole area of interest is larger than 10 000 km^2, for example, Landsat or similar orbital imagery would be the ideal starting point (given that it is available in useable condition, which is not the case at present for some parts of Africa). For much smaller areas, aerial photographs might be just as good if they are available and fairly recent.

One of the most important advantages of Landsat imagery over aerial photographs is that in the FCC format the Landsat imagery provides a geometrically corrected map of land cover, inexpensively for a large area on a single sheet. Most of the technical problems of analyzing and piecing together large numbers of aerial photos and making maps from the photos are already taken care of in the imagery product obtained from the EROS Data Center or one of the regional receiving stations. The land cover types, seen as combinations of colour, texture and pattern, can be relatively easily delineated on clear acetate overlays. Then the low-level aerial sample survey and/or local knowledge of officials can be used to label these delineated land cover types. The imagery, overlays and labels can be presented directly to decision makers with only brief orientation to the imagery and its interpretation.

The utility of Landsat and other orbital imagery for rural development depends on the size of the area in question and alternative sources of land resource information. If target areas have already been selected and are fairly small (for example, 500 to 2500 km^2), then Landsat is probably inappropriate. Traditional photo-interpretation of existing high to medium altitude, black-and-white aerial photographs is usually more economical, given that there is existing photo coverage of the target areas. The existing photos, however, may be too small in scale or the common black-and-white format may have insufficient contrast to show the required detail, and the photos may not be recent enough to show the current situation. In these cases, a project should acquire new aerial photographs designed to meet the particular needs of small area planning, as we did in Arusha Region.

Digital analysis of the Landsat imagery was seriously considered when planning this project, using the facilities of the Digital Image Analysis Laboratory at the University of Arizona. Two factors, however, persuaded us to rely on visual interpretation of photographic image products. The first concerned the cost, time and uncertainty of obtaining acceptable, deliverable products from digital analysis. While large area resource assessments have used digital analysis techniques, these techniques must be considered to be at a development stage, not yet operational. Our cost and time constraints dictated that we use well-tested operational techniques. The second factor was that the regional staff needed to participate in information interpretation to facilitate their use of the results in planning activities. The lack of facilities in Tanzania made participation in digital analysis difficult. Photographic image products were easily transported to and used in Arusha, and their interpretation was easily learned by the regional officials.

ACKNOWLEDGEMENTS

The work described here was carried out as part of the Arusha Planning and Village Development Project, funded by the Government of the United Republic of Tanzania and the U.S. Agency for International Development (USAID) and implemented by the Arusha Regional Development Directorate (RDD) and the U.S. firm, Development Alternatives, Inc (DAI), under USAID contract AID/afr-C-1556. We are grateful to Arusha RDD, DAI and USAID for their support throughout our work. Many individuals helped us in many ways; in particular, R Bryan, C Dodson, C Hillman, P Kitchens, R Mouat, D Peterson, T Peterson, C Sweet, and J Walton. D Cleveland, J de Jong, C Dodson, A Elias-Cesnik, C Hutchinson, P Mrocha, S Moore, M Parton, W Shaner, V Thomas, and E Whitehead contributed to the preparation of this paper.

REFERENCES

1 Arusha Region: Strategies and Priorities for the Year 2000, 1981. Regional Commissioner's Office, Arusha, Tanzania and Development Alternatives, Inc, Washington, DC.

2 Arusha Region Today: 1981, 1981. Regional Commissioner's Office, P O Box 3050, Arusha, Tanzania and Development Alternatives, Inc, Washington, DC 20001.

3 Bryan, R, 1980. Soil Degradation and Conservation in the Arusha Region. Development Alternatives, Inc. Washington DC.

4 Catalogue of Maps--Tanzania, 1980. Tanzania Division of Surveys and Mapping, P O Box 9201, Dar es Salaam, Tanzania.

5 Clegg, P. H, and J F Scherz, 1975. A comparison of 9 inch, 70 mm, and 35 mm cameras. Photogrammetric Engineering and Remote Sensing 41, pp 1487-1500.

6 Deriving Recommendation Domains for Central Province, Zambia, 1979. Cimmyt Eastern Africa Economics Programme, P O Box 25171, Nairobi, Kenya.

7 Dodson, C M, 1980. Land Information Materials Available to Arusha Planning and Village Development Project, Arusha, Tanzania. Development Alternatives, Inc, Washington, DC.

8 FAO Framework for Land Evaluation, 1976. Soils Bulletin 32, Food and Agriculture Organization, Roma.

9 Klingebiel, A A, and P H Montgomery, 1961. Land Capability Classification. Agricultural Handbook No 210, USDA Soil Conservation Service, Washington DC.

10 Livestock, Wildlife and Land Use Survey, Arusha Region, Tanzania, 1980. Ecosystems Ltd, P O Box 30239, Nairobi, Kenya. 3 vols.

11 Mackelwich, B E, *et al*, 1979. Information for Decisionmaking in Rural Development. Development Alternatives, Inc, Washington DC.

12 Mack, J G, 1980. Digital Processing of Remotely Sensed Images. National Aeronautics and Space Administration, Report SP-431, Washington DC.

13 Mouat, D A, J B Kepe, K E Pochop and B D Tunderwell, 1981. The use of remote sensing for an integrated inventory of a semi-arid area. Journal of Arid Environments 4, pp 169-179.

14 Norman, D W, 1980. The Farming Systems Approach: Relevance for the Small Farmer. Michigan State University Rural Development Paper No 5, Department of Agricultural Economics, MSU, East Lansing, MI 48824.

15 Norton-Griffiths, M, 1978. Counting Animals. Handbook No 1 (2nd ed) African Wildlife Leadership Foundation, P O Box 48177, Nairobi, Kenya.

16 Norton-Griffiths, M, T Hart and M Perton, 1982. Sample surveys from light aircraft combining visual observation and very large scale colour photography. University of Arizona Remote Sensing Newsletter 82-2, pp 1-4 (reprinted in ITC Journal 1983-1).

17 Peterson, D, and T Peterson, 1980. Range Management in Arusha Region: Discussion Paper. Development Alternatives, Inc, Washington DC.

18 Peterson, D, and T Peterson, 1980. The Village Profile Exercise' AFVDP Background Information, Impressions and Perceptions. Development Alternatives, Inc, Washington DC.

19 Planning Technologies Appropriate to Farmers: Concepts and Procedures, 1980. Cimmyt Economics Programme, Apartado Postale 6-641, Mexico 6, DF, Mexico.

20 Poulton, C E, *et al*, 1975. Potential Groundwater and Land Resource Analysis for Planning and Development. Earth Satellite Corporation, Washington DC.

21 Shaner, W W, P F Philipp and W R Schmehl, 1982. Farming Systems Research and Development: Guidelines for Developing Countries. Westview Press, Boulder, Colorado.

22 UNEP Selected Works on Ecological Monitoring of Arid Areas, 1980. GEMS PAC Information Series No 1, UN Environment Programme, P O Box 30552, Nairobi, Kenya.

23 USDA, Soil Survey Manual, 1951. Agricultural Handbook No 18, US Department of Agriculture Soil Conservation Service, Washington DC.

RÉSUMÉ

Des techniques de télédétection multi-niveaux ont été combinées pour fournir des informations sur les ressources et l'utilisation du sol pour le plan de développement rural dans la région d'Arusha en Tanzanie. Des images Landsat actuelles, aidées par des données d'enquêtes de postes des familles topographiques et les rapports existants sur la végétation et les conditions du sol ont été utilisées conjointement par des analystes d'image et des fonctionnaires de la direction agricole au niveau régional pour dériver les six secteurs de la région en unités de planification. Les fonctionnaires de la planification du secteur ont déterminé un maximum de cinq unités de planification pour des activités de planification et de développement prioritaires. Pour les régions prioritaires, les photos aériennes en couleurs naturelles ont fourni une information détaillée pour des discussions sur le plan d'utilisation du sol entre les fonctionnaires du district et les villageois. La considération de l'efficacité de cette approche par télédétection amène à recommander d'une façon générale son emploi pour des applications similaires. La technologie et le risque de temps du relevé des données et les activités d'interprétation devraient permettre un maximum de participation par les utilisateurs de l'information.

RESUMEN

Se combinaron técnicas de sensores remotos de nivel múltiple para obtener información sobre el uso y los recursos de suelo para su utilización en la planificación del desarrollo rural en región de Arusha, Tanzania. Utilizando conjuntamente imágenes de Landsat empleadas, complementadas con datos de surveys aéreos de bajo nivel, datos de ángulos de pendiente tomados de levantamientos topográficos y de los informes existentes sobre las condiciones de suelo y vegetación, los analistas de imágenes y los oficiales encargados de la administración de tierras distritales dividieron los seis distritos de la región en unidades de planificación territorial. Los oficiales de planificación distrital seleccionaron algunas de sus actividades para definir prioridades de planificación y actividades de desarrollo. Para las áreas prioritarias, la fotografía aérea de color natural aportó información detallada para las discusiones entre los oficiales distritales y los pobladores respecto a la planificación del uso del suelo. La consideración de la eficiencia de este enfoque de sensores remotos lleva a su recomendación general para aplicaciones similares. La tecnología y oportunidad del trabajo de recolección de datos y de la actividad de interpretación deberían permitir la máxima participación de los futuros usuarios de la información.

0777

A Microcomputer-Based General Photogrammetric System

P.-A. Gagnon, J.-P. Agnard, C. Nolette, and *M. Boulianne*
Faculté de foresterie et géodésie, Laval University, Québec, P.Q. G1K 7P4, Canada

ABSTRACT: A digital photogrammetric system, called a videoplotter, has been developed to perform photogrammetric tasks with a minimum of cost and complexity. With a microcomputer and split-screen imagery for stereomodel display, the problems of anterior, relative, and absolute orientation are solved in a rigorous and efficient way. With the possibility of storing orientation files with image files, the system can be used without specific training in photogrammetry, in a way similar to the use of other menu-driven software packages. In its present state of development, stereomodel processing with the videoplotter includes on-line compilation; determination of object coordinates; point, line, and surface superimpositioning in different colors; map updating; and digital terrain model automatic (DTM) generation. Accuracy, a function of input pixel size, is on the order of 0.05 mm at photograph scale, in the experiments presented. The use of a 450- or 600-pixel/inch digitizing scanner, now available at low cost, instead of a 300-pixel/inch one, would lead to a proportional increase of accuracy.

INTRODUCTION

IN A RECENT AND COMPREHENSIVE REVIEW of design concepts of digital photogrammetric systems, Gruen (1989) points out that photogrammetry has now fully realized the potential of the new working tools. And, although he states that we are still "at an early stage of development," it is apparent from the review that development activities are very intense and go in many directions. All the systems presented, however, because of their complexity and cost, cannot be considered directly accessible to the general photogrammetric community. This leaves open a line of research and development that is important to address and to which was directed the concept behind the system described in Agnard *et al.* (1988). This paper even stated that, because of the outstanding advances in microcomputer and microcomputer-related technology, photogrammetry now had the potential to be directly accessible-for reasons of costs, ease of operations, and flexibility-to a lot of new users in engineering, forestry, agronomy, architecture, archaeology, police investigations, etc. The idea was that the hardware was already available in the form of low cost standard equipment. What was needed at this point was the development of software capable of making optimal use of this equipment so that the system would solve the standard photogrammetric problems in a user-friendly and efficient way and even perform tasks that were beyond the possibilities of conventional photogrammetric instrumention.

This paper is concerned with the description and evaluation of a system which represents an example of materialization of this concept. The system constitutes the expected evolution of the much more limited prototype presented in Agnard *et al.* (1988) which was called the "videoplotter."

CHARACTERISTICS AND OPERATIONAL ASPECTS

The hardware components of the new videoplotter have been kept simple and standard. They consist of

- a Model 60 IBM-PS/2 with 1 MByte of RAM, a 40 MByte hard disk, an 80286 microprocessor, and a standard VGA card;
- a Model 8513 IBM Color Monitor with high resolution of 640 by 480 pixels (4 bits) and low resolution of 320 by 200 (8 bits); and
- a mirror stereoscope mounted in front of the monitor to facilitate stereoscopic vision.

Using digitized stereo-images previously stored on the disk, the operations follow the usual steps of interior, relative, and absolute orientations.

To facilitate, speed up, and make more flexible the interior orientation, the program, at the beginning of operations, brings to the screen a rectangular figure, similar in shape to each digital image, on which appears a color window that can be quickly positioned, with a cursor, at the approximate locations of the n fiducial marks ($n \geq 2$). Each part of the image containing a fiducial mark is then automatically brought to the screen, and a pointing to each mark is executed with the measuring dot while its calibrated coordinates are entered as input. Displacement to any place in the two displayed images can be achieved, at the operator's choice, by a combination of movements

- of one or the other or both images, or
- of one or the other or both superimposed measuring marks.

Photograph coordinates are determined by applying the standard orthogonal transformation to the pixel coordinates measured at the scanning stage and stored as a two-dimensional matrix. The transformation parameters are determined by applying a least-squares adjustment to the linear system of equations employing the calibrated and the observed coordinates of the fiducial marks. A different type of transformation could of course be implemented to meet different situations or more sophisticated needs. As directed by the menu, the operator introduces the focal length during the interior orientation procedure. When this operation is completed, the left and right images, centered respectively on the left principal point and its conjugate, are displayed as represented on Figure 1.

A decision was made to proceed through the intermediate step of relative orientation instead of going directly to the bundle formation because the procedure was better suited to some specific applications such as, for instance, the one described in Gagnon *et al.* (1989), where the photographs are considered to be vertical and the scale is obtained directly from the known distance between the cameras.

The relative orientation phase is accelerated by the fact that the software provides easy and fast movement to any part of the stereomodel. In the upper right corner of the screen, as is shown on Figure 1, two small rectangles represent the stereomodel and the black cursors, represent the position in the stereomodel of the displayed images. In order to have another part displayed, the white cursors are moved by a mouse to that part. Depressing a button then brings the black cursors to the same position and brings to the screen the images of that part. From there the measuring marks are moved to the desired conjugate points, by combining either mark or image displacements. Mark movement is usually preferred: it is independent from roam

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 56, No. 5, May 1990, p. 623-625

0099-1112/90/5605-623$02.25/0
©1990 American Society for Photogrammetry
and Remote Sensing

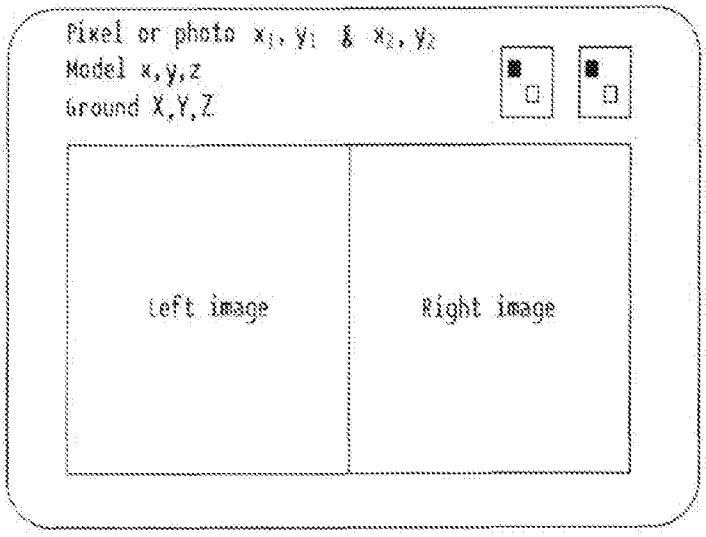


FIG. 1. View of the screen.

speed and probably seems more natural because it gives the impression of moving over the terrain instead of seeing the terrain move.

The relative orientation is solved for by least squares using the coplanarity condition, and a printout of the following is produced for analysis: the values of the orientation elements, the residuals y-parallaxes at the relative orientation points, and the standard y-parallax, at photograph scale. Once the relative orientation is solved, the measuring marks become locked to the respective conjugate points, so that the model is preserved in all displacement of the dots and/or of the images.

The ground or object coordinates $\mathbf{X}$ are computed from the model coordinates $\mathbf{x}$ by application of the similarity transformation

$$\mathbf{X} = k\mathbf{M}\mathbf{x} + \mathbf{X}^{\circ} \tag{1}$$

The scale factor k, the orthogonal matrix $\mathbf{M}$, and the vector of translations $\mathbf{X}^{\circ}$ are determined by least squares using the coordinates of the control points. The solution is based on the Space-M formulation, as described in Blais (1979). The seven independent unknowns of Equation 1 are computed by application of its following linearized form in an iterative way:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} a & b & -c \\ -b & a & -d \\ c & d & a \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} e \\ f \\ g \end{bmatrix} \tag{2}$$

Each iterative yields a new set of parameters $a, b, \ldots, g$, which are used to obtain improved values of k, $\mathbf{M}$, and $\mathbf{X}^{\circ}$ (no more than three iterations are sufficient to reach convergence). When the absolute orientation is completed, the system provides a listing of the control points, with their coordinates and discrepancies, and with the mean error in X, Y, and Z.

At the end of each of the three orientation steps, the operator has the possibility of storing the orientation observations or elements in a file attached to the image files, so that a given stereomodel needs to be oriented only once. For a subsequent use of the stereomodel, a menu-driven recall of these files will execute automatically the orientations, so that model processing can be started immediately. In the present state of development of the system, model processing includes

- on-line compilation, by linking the microcomputer to a cartographic plotter;
- determination of X, Y, Z coordinates in the object or terrain system. As shown on Figure 1, photograph, model, and terrain coordinates are continuously displayed at the top of the screen;
- point, line, and surface superimpositioning in different colors. As mentioned by Gruen (1989), this feature is useful for completeness checking. It is also used for map updating, data bank construction, and image analysis. In the color superimpositioning of surfaces, the relative densities of the covered images are preserved in the overlay, in the low-resolution option of the software, so that stereoscopic vision is maintained even inside the colored area; in the high-resolution option, the colors are flat (this limitation would disappear with the use of a card of the IBM-8514 type).
- DTM automatic generation. The stereomodel is scanned along lines parallel to the model x-axis, and terrain coordinates are computed at regular intervals on these lines. Each new DTM point is materialized on the screen by a green point superimposed on its conjugate corresponding points, which helps checking and keeping track of the operation as it proceeds. The point of best correlation is given by the maximum value of the following equation:

$$k = \frac{(\Sigma a_{ij} \cdot b_{ij})^2}{\Sigma a_{ij}^2 \cdot \Sigma b_{ij}^2} \tag{3}$$

where a_{ij} represents the grey value of pixel ij in the master (left) window and b_{ij}, the grey value of pixel ij in the slave (right) window. More sophisticated formulas have been tested but have not provided significantly better results. The result is subjected to three rejection tests:

- if k is smaller than a preset value,
- if a big jump in elevation is immediately followed by a similar jump in the other direction, and
- if the highest correlation comes from the extreme points of the slave (right) window.

When the computation cannot provide an adequate result, for instance, in the case of cliffs, high walls, or important shadows, a red point is displayed instead of a green one, as an editing tool, and a buzzer informs the operator.

SYSTEM PERFORMANCE

The last two features go beyond the possibilities of standard stereoplotters, conventional or analytic. The same applies in a certain way also to the second feature when considering the above-described possibility of going very fast to any point in the stereomodel, without having to take into account roaming speed, that characteristic contributing, in fact, to improving all the operations. The possibility of recalling and using the orientation files means that the system can be used without prior knowledge or background in photogrammetry, in a way that is similar to the use of other menu-driven software packages.

As the formulation of the different steps of operations rests on well-established mathematical concepts, and as pixel-size observations are easily performed on the type of screen used, the accuracy of the output is essentially ruled by the resolution of the input, which, in the experiments related to the development and testing of the system, has been generated by an HP Scan-Jet Digitizer, at a resolution of 300 pixels/inch, or 0.08 mm, in 16 shades of grey. Effectively, as the following results below indicate, the accuracy of the output is of the same order as the quality of the input. These results give a good indication of the present accuracy of the system; they were obtained using 1:6000-scale photography of the Canadian Research Council's test area of Sudbury, Ontario. The photography is 23 by 23 cm in format, taken with a 15 Uag objective Wild camera and has been digitized from paper positives. The model chosen contained 19 well-distributed X-Y-Z control points, known to the nearest 1 cm and materialized by 30- by 30-cm targets. Nine points were used as control points and ten as check points. The RMS, in metres, in X, Y, Z, was,

for control points:	0.23	0.32	0.08
for check points:	0.25	0.51	0.22

From these we get the following RMS, in mm, at photograph scale

for control points:	0.04	0.05	0.01
for check points:	0.04	0.09	0.04

With the same model, a standard y-parallax of 0.02 mm, at photograph scale, has resulted for the relative orientation, computed with ten well-distributed points, which agrees with results obtained with other models processed in a similar way. This result, in addition to the ones given above, is in accordance with the accuracy that can be expected from the input pixel size. For this reason, the scanner has only been calibrated for scale differences between its x and y axes, and the fiducial marks have been measured on the paper positives. The eventual use of a smaller-sized input pixel will, of course, require more sophisticated processing, in order to take advantage of the increased potential of accuracy. A substantial increase is already possible, considering the fact that low-price scanners are now available on the market, with resolutions of up to 600 pixels/inch instead of the resolution of 300 pixels/inch on which the above results are based. Another improvement will come with the possibility of digitizing the images with low-price color scanners. Interesting applications could be considered, in large-scale forest inventories, for instance, with a spread of intensity levels of colors sufficient to obtain displays that will appear reasonably smooth to the eye (Gonzalez and Wintz, 1987).

CONCLUSION

The systems described by Gruen *et al.* (1987) or Cogan *et al.* (1988), or the systems now being developed by Helava & Associates or Matra, have the possibilities of being very efficient and yielding very sophisticated and accurate results. Yet, as they are quite complex and costly, they are not readily accessible to users who, on the other hand, would be satisfied by the type of operations and the kind of results that the system described here can provide. This system meets the desirable requirements of flexibility, simplicity, and adaptability mentioned by El-Hakim (1985), at a hardware cost of under $5000, and much less if the user already has a suitable microcomputer at his disposal. Because of that, the category of potential users is vast: it includes not only those interested in the usual surveying and mapping applications but also a lot of new users, in remote sensing, forestry, agriculture, architecture, etc., who will most probably be able to see and develop numerous applications with this type of soft-copy photogrammetry.

REFERENCES

Agnard, J.P., P.A. Gagnon, and C. Nolette, 1988. Microcomputers and Photogrammetry. A new tool: the Videoplotter. *Photogrammetric Engineering and Remote Sensing*, Vol. 54, No. 8, pp. 1165–1167.

Blais, J.A.R., 1979. *Least-squares block adjustment of stereoscopic models and error analysis.* Doctoral thesis in the Department of Surveying Engineering, UNB, Technical report No. 30001, Division of Surveying Engineering, The University of Calgary.

Cogan, L., D. Cogan, D. Hunter, S. Lutz, and C. Peny, 1988. Kern DSP1 Digital Stereo Photogrammetric System. *International Society for Photogrammetry and Remote Sensing Congress*, Commission II, Kyoto, Japan.

El-Hakim, S.F., 1985. A Photogrammetric Vision System for Robots. *Photogrammetric Engineering and Remote Sensing*, Vol. 51, No. 5, pp. 545–552.

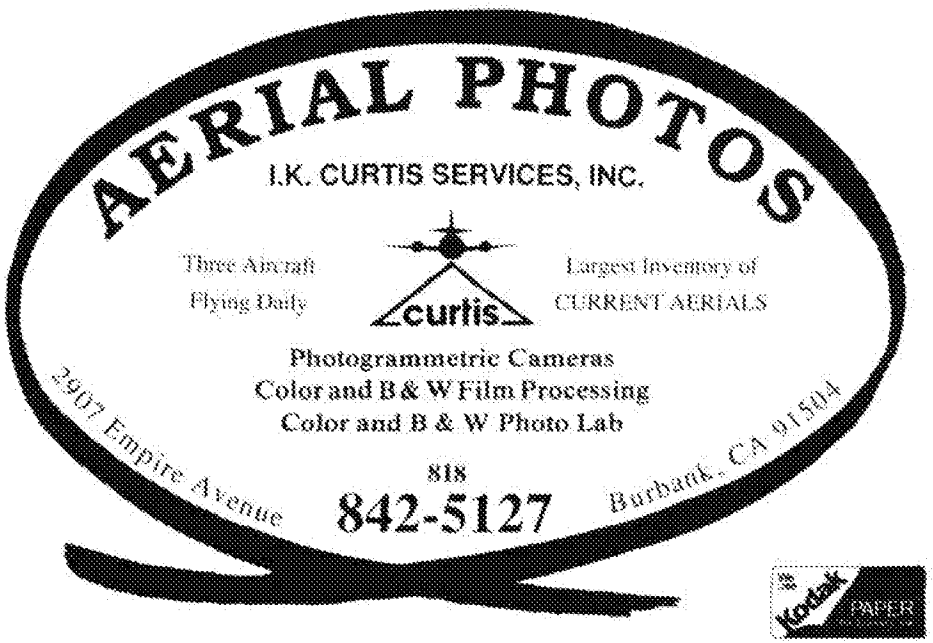
Gagnon, P.A., and J.P. Agnard, 1988. Twin-camera fixed-base photography of tree plots: possibilities and accuracy of a system. *Canadian Journal of Forestry Research*, Vol. 19, No. 7, pp. 860–864.

Gonzalez, R.C., and P. Wintz, 1987. *Digital Image Processing*, Second Edition. Addison-Wesley Publishing Co. 503 p.

Gruen, A.W., 1989. Digital Photogrammetric Processing Systems: Current Status and Prospects. *Photogrammetric Engineering and Remote Sensing*, Vol. 55, No. 5, pp. 581–586.

Gruen, A.W., and H.A. Beyer, 1987. Real-Time Photogrammetry at the Digital Photogrammetric Station (DIPS) of ETH Zurich. *The Canadian Surveyor*, Vol. 41, No. 2, pp. 181–199.

(Received 21 July 1989; accepted 15 September 1989)



Issues of Digital Mapping

by Gottfried Konecny, em. Prof., University of Hannover, Germany

Abstract

The traditional role of photogrammetric restitution has been fundamentally changed by the availability of new technical developments:

Digital cameras now compete with the established technology of analog aerial cameras. A new competitor is Pictometry and Multivision for gathering images by oblique images. The production of orthophotos as substitute for vector information has proven to be more economical. The most expensive part in orthophoto generation is the creation of digital elevation models. For DEM creation alternative methods have become possible, such as radar mapping and laser scanning. Finally vector data are now not aimed at CAD products, but as topologically structured data for use in a GIS. Photogrammetric hardware and software industry is currently in the process of providing direct shapefile extraction. The aim of photogrammetric data acquisition has changed from a map orientation toward the goal to provide GIS data.

Zusammenfassung

Die traditionelle Verfahrensweise der photogrammetrischen Auswertung zur Herstellung von Karteunterlagen hat sich durch neue technische Entwicklungen grundlegend verändert.

Digitale Kamerasysteme liegen heute in Konkurrenz zu der bewährten auf Film basierten Luftbildkameras. Neue Schrägaufnahmeverfahren, wie Pictometry und Multivision stellen die Notwendigkeit der bisherigen Luftaufnahme infrage.
Die Produktion von Orthophotos als Strichkartenersatz ist weit wirtschaftlicher als die Herstellung von Vektorinformationen. Der teuerste Aspekt bei der Herstellung von Orthophotos ist die photogrammetrische Erfassung digitaler Höhenmodelle. Dafür kommen aber auch Alternativmethoden, wie die Radarablastung und das Laserscanning infrage. Schliesslich sind neuerdings die Vektorinformationen in Form von topologisch strukturierten Daten für eine Verwendung im GIS gefragt. Die photogrammetrische geräte- und software herstellende Industrie ist gegenwärtig dabei die direkte Shapefile- Datenerfassung zu realisieren. Das Tiel der photogrammetrischen Datenerfassung hat sich von der Strichkartenorientierung zur Bereitstellung von GIS- Daten gewandelt.

1. Introduction

Mapping by traditional field surveys has been replaced in the 20th century using tools of aerial photogrammetry, which is now an efficiently operating, highly automated technology. But more recently during the past decade new technological aspects

OFEK07601

have been introduced to improve the mapping technology even further. Some of these trends are discussed in this paper. (see Konecny 2003).

2. Digital Cameras versus Scanned Film Camera Images

Ever since the progress of U.S. military and that of the U.S. space program in the 1960´s and 1970´s digital cameras have been used to transmit images from orbits. More recently the consumer market has made a breakthrough for 5 Mpixel and more digital amateur cameras.

To compete with the high level requirements of film based aerial survey cameras with one highly optimized wide angle objective a 900 M pixel (30 000 x 30 000) camera would be required, which is still too costly to realize, unless shortcuts are taken.

There is the option to use several lenses for imaging onto smaller CCD area array chips. This option has been realized in the Z/I Intergraph DMC with 7 lenses and the Vexcel Ultracam.

Another option is to develop the camera in form of a CCD line scanner. This option, used by satellite systems, was chosen for the Leica ADS 40 camera.

The technical details of these commercial and operational camera systems are given in table 1. (see Graham, Hinz 1999, Leberl 2005, Trinder 2005).

Table 1: Parameters of Operational Digital Aerial Cameras

Type	no.of pixels	pixel size	ground coverage at h=1000m	h/b	$\sigma_{x,y}$	σ_z
Vexcel Ultracam D	7500 x 11 500	9µm	675 x 1035 m at f=100mm	3.7	+/-9cm	+/- 16.6cm
Z/I Intergraph DMC	7600 x 13 824	12µm	768 x 1383.4 m at f=120mm	3.3	+/-10cm	+/- 16.5cm
Leica ADS 40	12 000 lateral longitudinal: 28.4° forward nadir 14.2° aft	6.5µm staggered	swath 1924m at h=1540m (15cm ground pixels)	1.26	+/-15cm	+/- 20cm

The imaged areas by the multilense area chips have to be resampled and fitted together according to boresight calibration information by subsequent image processing. The fitted image corresponds to resampled area arrays for a usable photogrammetric image of 7600 (longitudinal) x 13 824 (lateral) pixels for the DMC and of 7500 (longitudinal) x 11 500 (lateral) pixels for the Ultracam. Even forward motion compensation is included for the resulting image.

The radiometric resolution of digital sensor elements is by far superior (11bit grey levels) to that of the scanned film, for which due to exposure conditions usually only 6 bit of grey level information can be retrieved. For a particular exposure setting.

The same holds true for linear camera arrays. Nevertheless the resultant scanner images need to be corrected for aircraft forward motion and aircraft attitude, which are recorded by a rather expensive IMU system, which has to register the 6 orientation components in real time during the flight. Linear array camera systems therefore depend during their operation on the speed of the aircraft and its attitude changes, seeting a limit to the achievable ground sampling distance, depending on the frequency at which the IMU records can be collected. For that reason the ADS 40 May have difficulties to achieve the same ground sample distance for low level flights, as obtainable for cameras using area chips exposed at a specific time interval.

The standard flight planning procedures for analog aerial cameras need of course modifications when digital cameras are used, since the coverage on the ground depends on the chip sizes and the different focal lengths utilized for them.

The other issue is the achievable accuracy on the ground for position and elevation. Accuracy is of course a function of the ground sample size and the base-height ratio, depending not only on the the lense opening angle, but also on the method of data extraction (automatic or manual). No doubt extensive accuracy tests will be made in the coming years by independent bodies to clarify the accuracy issue.

Another aspect is the product obtained by digital cameras: While cameras with area arrays, after boresight calibration and preprocessing can be treated in almost the same way as analog images for aerial triangulation, DEM extraction, orthophoto generation and line mapping, the line scanner cameras more clearly depend on the processing algorithms of the manufacturers, who of course are ready to license the appropriate software packages for processing.

The digital camera technology using Canon, Kodak or Rolleimetric cameras do not quite reach the resolution and geometric accuracy requirements of the DMC, the Ultracam and the ADS 40, since they have not been specifically designed for photogrammetric mapping. Nevertheless Canon, Kodak and Rolleimetric cameras have been successfully used for mapping applications in developing countries at reduced geometric requirements. Since these cameras can be accommodated in small aircraft (Cessna, Piper) their use is possible in remote areas, for which the mobilization of a special survey aircraft is too costly.

Such cameras can, on the other hand, also be combined into cuch camera assemblies have been used camera assemblies using vertical and oblique images in up to 4 directions. Such camera assemblies have been used by the Rochester, N.Y.

based U.S. company Pictometry. The use of the Pictometry system is to enable views of all house facades of a city from oblique images and to use the 4 oblique images for linear measurements on the imaged objects (e.g. distance between two houses). The Pictometry camera systems are used from 2 flying heights producing higher resolution images from low level flights for the city core areas in form of "Neighbourhood Images", and from high level flights for the outskirts of a city in form of "Community Images" at a lower resolution. Extensive use of this technology has been made in the USA in the "Homeland Security Program". (see fig. 1, 2 and 3 and www.pictometry.com)

In Israel Ofek-Air has produced a similar software system "Multivision", in which conventional orthophotos can be viewed and used for measurements together with georeferenced digital oblique images. (see fig.4 and 5 and www.ofek-air.com).

Pictometry and Multivision offer the possibility to view and measure object dtails without the need for conventional mapping. They are therefore a serious competition in security and emergency applications.

3. Orthophoto Mapping versus Line Mapping

Traditionally line mapping was the target of the photogrammetric mapping process. Ever since the introduction of the orthophotorectification methodology orthophotos have been seen as map substitutes, which were faster and less expensive to produce. While a line mapping project may take 2 to 3 years to complete, the orthophotos for the same area may be completed in 6 months. Particularly in developing countries with a fast growing urban population and a lack of systematic updating systems for maps, orthophotos offer the only possibility to produce quick information for urgent needs.

Of course the cost aspect is even more important, since urban centers in developing countries can seldom afford a line mapping effort.

The costs for aerial mapping are based on the number of photos (or stereomodels) required to cover a specific area. A summary of costs of the components of an aerial photogrammetric project is given in table 2.

Table 2:
Orthophoto Mapping versus Line Mapping Costs (international average)

Process	Remarks	Cost per Aerial Photo
aerial photography mobilization	constant fee per block $5000	-
aerial photo		10 $
scanning of photos		15 $
aerial triangulation and block adjustment		25 $

digital elevation model		120 $
digital orthophoto production		30 $
mosaicking of orthophotos		20 $
on screen digitization in 2D or 3D	rural area 10^b per image @ 50$/hr	500 $
	urban area 100^h per image	5000 $

Mobilization for an aerial photography project is based on a fixed cost depending on the accessibility of an area, which can be high for a remote area lacking a mobilization infrastructure for operating a survey aircraft.

Scanning of photos is not required for use of digital cameras, while the costs per image, aerial triangulation, the generation of a digital elevation model, the orthophoto production and the mosaicking of orthophotos are similar for analog and digital images. However, the availability of digital cameras is still limited due to their high investment cost exceeding 1 M $.

The costs for generation of line maps using 2D procedures (for rural areas) and using 3D procedures (for urban areas) are quite high, depending on labour cost. Many mapping consortia in Europe and North America have therefore subcontracted the line mapping tasks to countries, where labour costs are still low (e.g. Eastern Europe, India and China).

Table 3 and table 4 compare the total cost for an urban aerial mapping project covering 250 km^2 according to average international prices.

Table 3:
Costs for a City Map covering 250 km^2 with 20 cm ground pixels, scanned with 15µm from photography at image scale 1:13 333:

Process	**no.of photos**	**cost**
aerial photography	83 photos	$ 4830
scanning		$ 1245
aerial triangulation		$ 2075
DEM		$ 9960
digital orthophoto		$ 2490
mosaicking		$ 1660
total orthophoto cost		**$ 22 260**
line map digitizing		$ 124 500

total line map cost		**$ 146 760**

Table 4:
Costs for a City Map covering 250 km^2 with 10 cm ground pixels, scanned with 15µm from photography at image scale 1: 6666:

Process	no.of photos	cost
aerial photography	332 photos	$ 7 320
scanning		$ 4 980
aerial triangulation		$ 8 300
DEM		$ 39 840
digital orthophotos		$ 9 960
matching of orthophotos		$ 6 640
total orthophoto cost		**$ 77 040**
line map digitizing		$ 498 000
total line map cost		**$ 575 040**

The total bidding prices for a mapping contract still depend on the following factors:

- production costs (as seen in tables 3 and 4)
- field verification add 30 %
- overhead add 20 %
- profit add 10 %
- risks add 40 %

This is why a bid may be based on a cost plus 100 % price. This may of course vary on the competitive economic situation.

When comparing pricing for a city map **1: 2000** produced from 20 cm ground pixels, this **line map** may be contracted for **$ 293 520**.
The **orthophoto** coverage would cost **$ 44 420**

A city map **1: 1000** produced from 10 cm ground pixels
would be contracted for **$ 1 150 080**
and the **orthophoto** for **$ 154 080**

This compares to an alternate **ground survey** procedure using GPS/DGPS and total stations at $ 6400 per km^2 for **$ 1 600 000**.

Photogrammetric mapping is still less costly than ground surveys, and orthophoto production is the most affordable procedure.

Particularly in urban areas orthophotos pose the difficulty, that the geometry of the orthophoto is based on ground elevations. Therefore the tops of buildings or bridges are displaced in the orthophoto. A solution to this problem may be the production of socalled "True Orthophotos", in which the tops of the buildings or objects, produced by a digital surface model, may be shifted to obtain the true orthophoto. This, however requires manual intervention, which is costly. The objects to be shifted must

first be digitized in 3D. E.g. the company Inpho has effective software to minimize the effort in dong this. But still the costs for producing a true orthophoto may be 2 or 3 times as high as for the standard orthophoto product. (see Schickler 2003).

The clients, deciding on a particular product strategy should therefore seriously consider the questions:

- what ground sample distance is ultimately needed under financial constraints (e.g. utility manholes may not be visible with 20 cm ground pixels, but utility features may be more easily determined using GPS/DGPS ground surveys. For the building topography a 20 cm ground pixel may be sufficient).

- is it necessary to rectify roof top levels to be able to refer the orthophoto to cadastral (and most often not updated) boundary information.

4. Digital Elevation Models and Laser Scanning

From the cost comparison of photogrammetric processes it becomes clear, that one of the high cost factors is the generation of a digital elevation model by photogrammetric means.(either image matching, semiautomatic grid DEM measurement or manual contouring). If a digital elevation model is already available (presumably it does not change very often) the costs for orthophoto production may therefore be significantly reduced.

This is the case, if an older contour map is digitised and interpolated into a rectangular DEM grid, which can be further interpolated for orthophoto production.

There are also alternative methods available to generate digital elevation models and digital surface models.

As has recently been shown by the Intermap Co. from Calgary, Alta., Canada a digital elevation model with +/- 0.5 m accuracy has been generated for one third of the area of Britain (50 000 km^2) and with +/- 1 m accuracy for the additional area of two thirds of Britain (100 000 km^2) with 5 m posting at a cost of 6 $/$km^2$ per customer. With the Star-3i airborne radar in the "Nextmap Britain" project.

Satellite radar systems, such as those flown on the Space Shuttle Radar Topographic Mission (SRTM) provided a nearly worldwide DEM at 90 m posting with +/- 5 m accuracy in areas not affected by radar shadows and layover in mountain areas. For the U.S. the posting of the data has been reduced to 30 m. Even though the posting and the accuracy is still rather low, the information may be utilized with near vertical high resolution satellite images (Quickbird, Ikonos) to produce digital satellite image orthophotos of 2 to 3m accuracy. Also Google Earth makes use of the information for viewing. (see Henderson 1998).

Another technology is airborne LIDAR. At present about 140 Lidar systems operate worldwide (most of them are made by Optech, Leica and TopEye). Newer Lidar systems operating at 100 KHz pulses are able to generate digital surface models (first pulse return), digital elevation models (last pulse return) and intensity images. Depending on the flying height ranging between 100m and 6000m sub-meter to 10 m

postings may be generated with dm to m accuracy. (see Ackermann 1999 and Baltsavias 1999).

The problems with Lidar data are caused by the fact, that the GPS/DGPS CORS and the IMU data are generally not timed accurately enough with respect to the more rapid laser pulses. Furthermore the identification of ground points in the laser image is not as accurate as in optical images.

Therefore cross flights must be used to mosaic the laser images. To geocode them the use of elevated targets are the best practical solution. In the future the simultaneous operation combined with optical digital cameras is expected.

5. CAD Mapping versus GIS Shapefile Mapping

Most digital mapping efforts have in the past served automated map production in form of Autocad or Microstation files. With these a great number of feature codes have been used with annotated symbology added to the graphics, which does not have intelligence.

As mapping must now be considered as a base for the analysis capabilities of a GIS, there is the need to adapt the photogrammetric data acquisition into an object oriented data structure. As GIS shapefiles carry the data topology and give access to attribute data entries into object relational databases, photogrammetric data acquisition should use these advantages. (see fig. 6 and 7).

For new data acquisition several hardware/software manufacturers now provide or attempt to provide practical solutions for direct shapefile creation during the 3D digitisation process in digital photogrammetric work stations:

- Inpho-DATM
- KLT
- Leica LPS
- Socet Set 3D
- Z/I Intergraph

As many mapping companies already own efficient workstation software, e.g. the Microstation based Z/I Image station software, or the use of Bentley´s Geographics, the alternative to the direct solution is the conversion of Microstation or Autocad data into shape files, even though this is not a fully automated process.

While the topology required can be generated automatically into shapefiles for most of the features, the intelligence must be entered manually into attributes attached to label points.

This will help to reduce the required number of feature codes, which in any case are impractical to update.

This semiautomatic conversion procedure is in any case required for already collected datasets in Microstation or Autocad files.

The GIS concept, as demonstrated in the ESRI ArcGIS platform, clearly aims at the generation of an updatable object oriented geodatabase solution. Photogrammetric Operations must now be aimed toward this goal.

The advantage of a proper geodatabase design lies in the possibility to separate point coordinate information for specific points (e.g. boundary points) from topologically related line and area information. This permits to generate the topology of land parcels, buildings and other objects, regardless of the accuracy with which the nodes (boundary points) have been determined. It is therefore possible to use rather low resolution ortho-images with ground pixel sizes ranging from 20 cm to 2 m to create the object topology first, and to improve the point coordinates in a sporadic manner by DGPS ground surveys, when the need arises in transactions.

The object relational database attributes attached to points, line segments, areas and defined objects are the basis for GIS analysis.

6. Conclusions

The present state of technological development in photogrammetry is marked by a transition from an established, but isolated and self sufficient mapping technology into a GIS integration concept. The product is not a map any more, but base data within a GIS.

Analog aerial camera technology is gradually replaced by digital camera technology.

Cost effective GIS products, such as orthophotos contained in a geodatabase in raster form are combined with vector data in superposition. A combination of survey methods ranging from terrestrial DGPS uses for certain features to 3D photogrammetric data acquisition and laser scanning offers a cost optimise approach for new mapping or its updating.

Finally, the GIS data base concepts should be the guidelines along which photogrammetric products should be developed.

Photogrammetry is presented with anew challenge by these developments.

7. Literature

Ackermann, Friedrich: "Airborne Laser Scanning, present status and future Expectations"
ISPRS Journal of Photogrammetry and Remote Sensing
54 (2-3) 64-67, 1999

Baltsavias, Emanuel: "Airborne Laser Scanning: basic relations and formulas"
ISPRS Journal of Photogrammetry and Remote Sensing
54 (2-3) 199-214, 1999

ESRI "Shapefile Technical Description"
ESRI White Paper, July 1998

OFEK07609

Graham, Ron & Koh, Alexander:
"Digital Aerial Survey, Theory and Practice
ISBN 1-870325-98-2

Henderson, Floyd M & Lewis, Anthony J.
"Principles and Applications of Imaging Radar"
Manual of Remote Sensing, Third Edition, Vol. 2
John Wiley & Sons, New York 1998
ISBN 0-471-29496-3

Hinz, Alexander: "The Z/I Digital Aerial Camera System"
Photogrammetric Week '99'
Wichmann Verlag, Heidelberg, 1999

Konecny, Gottfried: "Geoinformation"
Taylor & Francis, London 2003
ISBN 04-15237955

Leberl, Franz & Gruber, Michael:
"Ultra Cam D – understanding some noteworthy capabilities"
Photogrammertic Week 2005
Wichmann Verlag, Heidelberg, 2005

Schickler, Wolfgang & Thorpe, Anthony
„Operational Procedure for Automatic True Orthophoto Generation"
IAPRS Vol. 32/4
ISPRS Comm. IV Symposium on GIS, 2002

Trinder, John: "Characteristics of New Generation Digital Aerial Cameras"
Map Middle East Conference 2005
Conference Proceedings

Photogrammetria – Elsevier Publishing Company, Amsterdam – Printed in The Netherlands

ANALYTICAL AERIAL TRIANGULATION WITH CONVERGENT PHOTOGRAPHY [1]

G. KONECNY

Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B. (Canada)

(Received June 28, 1966)

SUMMARY

Some of the inherent weaknesses of convergent analogue methods can be improved by a more rigorous analytical approach. Systematic interior-orientation errors can be eliminated analytically by applying valid lens distortion corrections. Nine locations can be used for the determination of relative orientation, the zero-base model can be dealt with, and triplets and quadruplets can be analyzed simultaneously to restrict error propagation. A favorable different selection of rotation axes can be selected. Special condition equations are shown for formulating a rigorous computational adjustment.

INTRODUCTION

Since the advent of superwide-angle photographs studies about convergent photography have decreased considerably in economic importance. Convergent photography, regardless of the angular field of the camera, however, remains a photographic system by which still the largest area may be covered at a given photo scale, representing a corresponding resolution and point accuracy. The treatment of convergent photographs also remains of academic interest. It is related to the problem of evaluating multidirectional photography from an exposure station in space which may have future practical significance. Furthermore convergent photography represents a study object for which many principles of general nature may be explored better than for simplified vertical photography.

Investigations about convergent photographs and specifically about aerial triangulation with convergent photography were sofar almost exclusively limited to the use of analogue plotters. Aerial triangulation tests were first made by PENNINGTON (1954), THEIS (1958) and GRIFFIN (1959). They did not yield a

[1] Paper presented at the International Symposium on Spatial Aerotriangulation, February 28–March 4, 1966, University of Illinois, Urbana, Ill. (U.S.A.).

sizeable accuracy increase over vertical procedures. This was contrary to theoretical predictions by VON GRUBER (1930) and HALLERT (1954). The author (KONECNY, 1962) demonstrated theoretically and practically that significantly more accurate results, comparable to theoretical predictions can be obtained if errors of interior orientation are taken into account or can be eliminated.

Other contributions along theoretical lines were made by ACKERMANN (1956) who derived numerical relative orientation formulae for overlapping stereophotographs and for simultaneous exposure-station photographs on analogue plotters. THOMPSON (1957) and RINNER (1957) made the first suggestions for an *analytical* treatment of convergent photographs.

The work of ROSENFIELD (1957) and ARTHUR (1956) on analytical relative orientation of photographs taken from the same exposure station is also directly related to the analytical treatment of convergent photography.

The suggested analytical formulations have sofar followed an analogue approach. This paper makes an attempt to improve the weaknesses inherent to convergent analogue methods by a more rigorous analytical approach. Some of the problems associated with aerial triangulation of convergent photographs are first to be discussed.

PROBLEMS ASSOCIATED WITH AERIAL TRIANGULATION OF CONVERGENT PHOTOGRAPHS

Interior orientation

In a previous article the author (KONECNY, 1965) studied the effects of interior orientation errors on photogrammetric accuracy for convergent photographs. It was shown, that the model accuracy was strongly influenced by errors of interior orientation. They caused additional y parallaxes in the model. This, in turn, resulted in an erroneous relative orientation. Fig.1 illustrates model deformations of the particular orientation elements which act differently in convergent photographs.

Systematic errors of interior orientation such as those of camera and instrument calibration cause a more noticeable propagation of systematic errors in the strip. Accidental errors of interior orientation, such as those of centering the diapositive plates in the plate holders also propagate systematically through double summation of relative orientation errors.

The overall accuracy of triangulation thus suffers from the propagation of the larger relative orientation errors, which result from a combination relative and interior orientation elements.

Zero-base model and error propagation

One further disadvantage of convergent photography for aerial triangulation

0792

OFEK07661

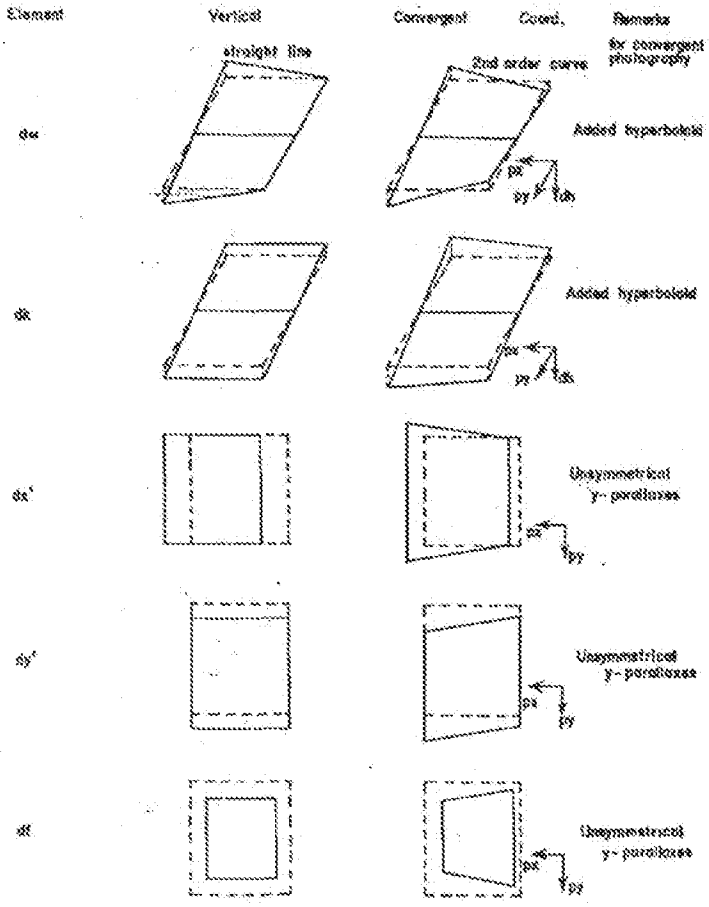


Fig.1. Comparison of model deformations for vertical and convergent photography.

is its property that one convergent photograph can only be used for one specific stereoscopic model, while a vertical photograph can be utilized in two stereomodels. Convergent models have therefore the advantage of a square sized model of 90–100% longitudinal overlap (see Fig.2). But the connection from model to model must be made in a narrow overlap from practically the same exposure station. The relative orientation of the synchronized cameras due to lack of sufficient overlap in the narrow "stereopontometer-" or "zero-base model" has to be made eliminating x and y parallaxes. Height displacements should not become noticeable in any case, due to the negligible separation of the nodal points of both cameras of a few inches and due to rigid synchronization specifications of both camera shutters.

This relative orientation of the "zero-base model", which is also subject to interior orientation errors accounts for $(2n - 1)$ different relative orientations to

OFEK07662

be performed in order to triangulate n stereomodels. Theoretical considerations show that error propagation is up to six times larger for convergent as compared to vertical photographs. In a practical attempt to relate point accuracy $m_{p,h}$ after triangulation adjustment in a strip with the numbers of models bridged n, BRANDENBERGER (1957) quotes a rough empirical relationship $m_{p,h} = a \times n^b$ in which a is a constant and b is a coefficient describing the error propagation.

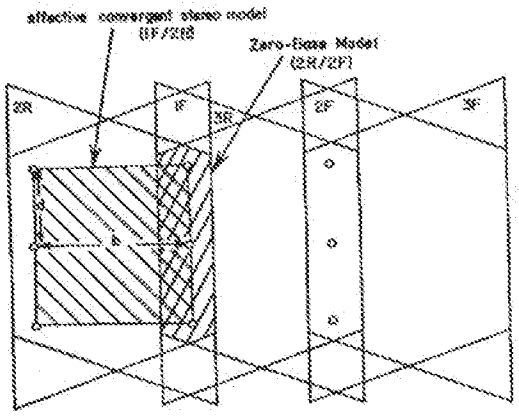


Fig.2. Convergent overlaps.

For vertical photography b is in the order of 0.5. Tests with normal angle convergent photography by the author resulted in a value of b close to 1.5 (see

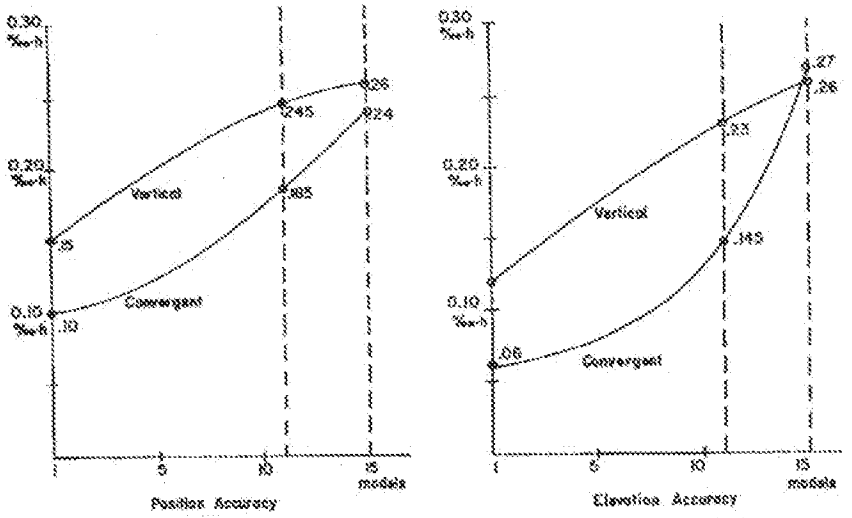


Fig.3. Comparison of accuracy of aerial triangulation after strip adjustment for simultaneous convergent normal-angle and vertical-wide-angle photography.

OFEK07663

Fig.3). Even though this comparison constitutes nothing but a thumb rule, it illustrates that the error propagation of convergent photographs is worse. While the increased accuracy of convergent procedures remains effective for short strips of up to fifteen models, analogue aerial triangulation of long strips shows no advantages.

OBJECTIVES OF AN ANALYTICAL APPROACH

Some of the difficulties encountered with analogue aerial triangulation of convergent photographs may be avoided using analytical procedures.

Interior-orientation errors

Systematic interior-orientation errors of the point of symmetry of the photographic camera may easily be eliminated analytically by precalibrated lens-distortion corrections. Interior-orientation errors practically do not occur in comparator measurements since the fiducial marks of the diapositive are read more accurately than it is possible to center diapositives in a stereo plotter. Interior-orientation errors therefore become ineffective for analytical procedures.

For relative orientation of convergent photography of both the stereo and the zero-base model it also becomes possible to eliminate interior-orientation errors dx'' and dy'' out of image-coordinate measurements if points are measured at nine characteristic locations (see Fig.4), giving order and rank of seven for the orientation matrix consisting of seven linear equations.

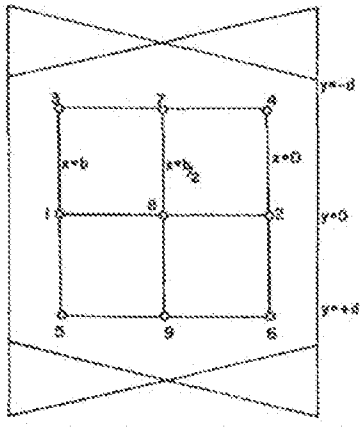


Fig.4. Relative orientation points at nine characteristic locations.

For vertical photography for an order of the orientation matrix of seven the rank of the matrix would remain five since dx' and dx'' as well as dy' and dy'' are

OFEK07664

linearly dependent. Thus dx'' and dy'' could not be determined for vertical photography unless dx and dy was accurately known. The terms dx' dy' can be considered known from the previous relative orientation except for the first model. Elements df' and df'' need not be included since they are related to film-shrinkage compensation. Considering that the variances in the variance-covariance matrix of the relative orientation system for dx'' and dy'' are rather large as compared to the variances of the other orientation elements the inclusion of interior orientation elements into a relative orientation solution is not advisable unless the camera parameters are entirely unknown.

Error propagation

While analogue approaches require the orientation of models in succession: stereomodel—zero-base model, stereomodel, etc.—it is possible to utilize redun-

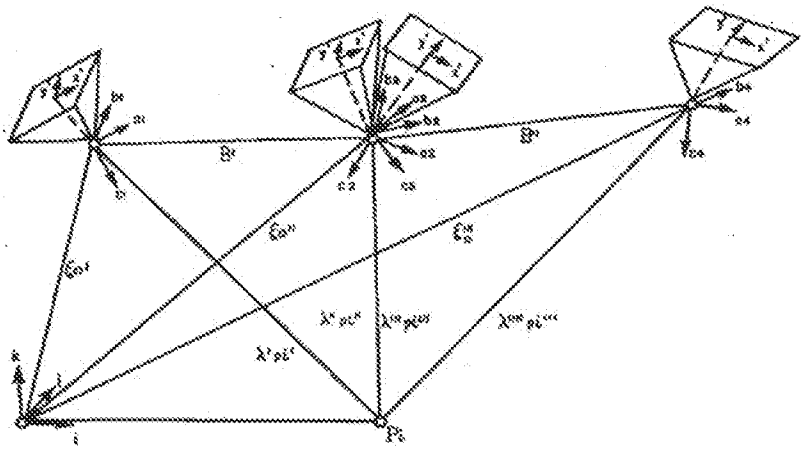


Fig.5. Convergent "quadruplet" with identical twin exposure stations.

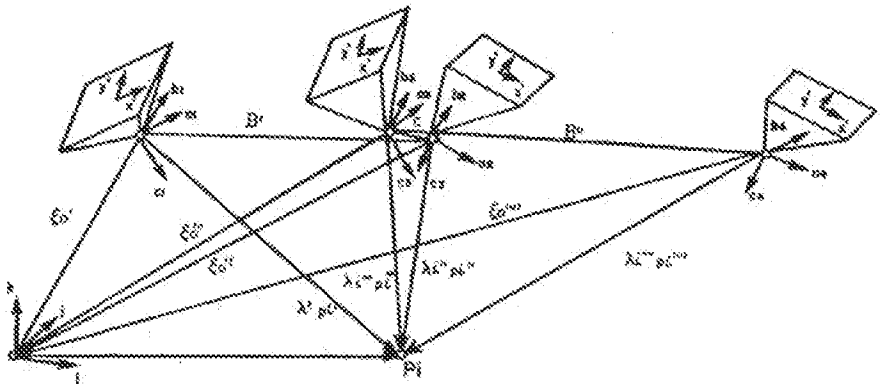


Fig.6. Convergent "quadruplet" with separate twin exposure stations.

0796

OFEK07665

dant geometrical relationships from additional photographs to permit a more accurate simultaneous relative orientation of multiple cameras.

In vertical photography there is a theoretical accuracy increase of 20% in triplets over the classical model-orientation method (LEHMANN, 1963). Triplets utilize, wherever possible, three rays instead of two to simultaneously compute the relative orientation elements of the three cameras. Classical model orientation neglects the existing correlation and encounters a loss in accuracy due to a stronger error-propagation possibility.

In convergent photography four rays intersecting at a point are being made use of instead of two in a configuration which may be called a "quadruplet" (see Fig.5 and 6). Additional angular conditions can easily be enforced analytically. A higher increase of accuracy due to diminished error propagation will therefore result for convergent photographs than for triplets.

Instrumentation

One of the drawbacks in triangulating and plotting convergent photographs has been the limitation of certain analogue plotters to accommodate this type of photography and to view it without undue strain.

Experiences in the Jena 1818 Stereocomparator and in the Wild PUG-3 point transfer device with normal-angle convergent photography taken with a twin camera Zeiss RMK $2 \times 21/18$ have shown no major difficulties to measure coordinates stereoscopically or to transfer points for use in a monocomparator.

MODIFICATIONS OF EXISTING COMPUTER PROGRAMS

When evaluating a strip of convergent photography by analytical means for the first time one prefers to find a simple workable solution which may perhaps not be the most suitable, but which will quickly yield a practical result without major programming effort.

Following this approach the National Research Council Aerial Triangulation Program by SCHUT (1964) was modified for specific use with convergent photography on an I.B.M. 1620 electronic computer.

Only the relative orientations of individual models of the Schut program could be used directly. In doing so the existing correlation between measured photocoordinates x' and y' in the coplanarity equations had to be neglected. (For vertical photography only the y' coordinate measurements contribute to relative orientation since the coefficients of the dx' or vx' terms become zero. Thus the coplanarity equations can be adjusted for vertical photography as simple observation equations.)

The program converges well for the orientation of convergent photographs however and constitutes the first subroutine of the triangulation program.

OFEK07666

For the triangulation of a strip two other supplementary subroutines were written in FORTRAN language: (*1*) a program for the orientation of the zero-base model and (*2*) a routine to calculate the scale factors λ_1 and λ_2 in two adjacent models for a vector from the exposure station common to both models to a ground point. Their ratio is used to obtain corrected values for *bx*, *by*/*bx* and *bz*/*bx* for the subsequent model.

The three subroutines applied in succession permit to triangulate a convergent strip according to analogue instrument principles. Interior-orientation errors are minimized by the analytical approach, but additional conditions minimizing the error propagation are not made use of. Therefore a new program was designed for use in a small computer, such as the I.B.M. 1620.

THE NEW PROGRAM

Choice of formulation

Before the decision was made, as to which formulation should be adopted for the analytical treatment of convergent photographs in a small computer, a comparison was made between formulations utilizing coplanarity equations according to SCHUT (1964) and those using collinearity equations according to SCHMID (1956). Schmid's formulation has the advantage that for the usual assumption of correlation free measurement of photo coordinates *x'* and *y'*, the correlation matrix reduces to unity also for convergent photographs. There is however the alternative of adjusting ground coordinates simultaneously with the relative orientation elements, as this has been suggested by SCHMID (1956), or to introduce absolute changes of these ground or model coordinates in terms of relative orientation elements as this has been done by HARRIS et al. (1962). The latter approach leads to unnecessary complications for more than two camera stations. The former is less suitable for small computers since it increases the number of unknowns to be simultaneously solved for. When partitioning methods are applied this is not necessarily a limiting factor for small computers, but the use of coplanarity equations is simpler. Due to limited computer size a subsequent strip and block adjustment to ground coordinates will in any case be necessary at a later stage.

Choice of rotations

A further look was taken at the choice of rotational elements to be used in the coplanarity equations. The choice of rotated axes α_1, α_2, α_3 applied in succession, such as those introduced by SCHUT (1964), is preferable to customary fixed sequential axes φ, ω, $\varkappa$ on account of the ease of derivation as well as calculation of coefficients. The use of rotated axes for convergent photographs, however,

0798 OFEK07667

requires that appropriate φ values are introduced in the calculation for the elements of the approximate orientation matrix of each photo; otherwise convergence of the solution is too slow.

Existing types of conditions

The well-known conditions of relative orientation and scaling of vertical photography are also valid:

(*1*) The *relative orientation condition* in form of the coplanarity equation:

$$\begin{vmatrix} bx' & by' & bz' \\ u' & v' & w' \\ u'' & v'' & w'' \end{vmatrix} = 0$$

(*2*) The *scaling condition* in form of the equation expressing the intersection of three corresponding rays in one point:

$$\begin{vmatrix} bx' & bz' \\ u' & w' \end{vmatrix} \begin{vmatrix} u'' & w'' \\ u''' & w''' \end{vmatrix} + \begin{vmatrix} bx'' & bz'' \\ u''' & w''' \end{vmatrix} \begin{vmatrix} u'' & w'' \\ u' & w' \end{vmatrix} = 0$$

in which u' v' w' represent rectified photocoordinates x' y' and f; the rectification is made by a rotational matrix $\mathbf{A}'$ with the elements a_1' to c_3' so that:

$$\begin{pmatrix} u' \\ v' \\ w' \end{pmatrix} = \begin{pmatrix} a_1' & b_1' & c_1' \\ a_2' & b_2' & c_2' \\ a_3' & b_3' & c_3' \end{pmatrix} \begin{pmatrix} x' \\ y' \\ f' \end{pmatrix}$$

A 1-prime indicates elements of camera 1, a two-prime of camera 2, and a 3-prime of camera 3. bx', by' and bz' represent the base components of model 1 (bx'', by'' and bz'' of model 2).

Furthermore the following conditions are valid for convergent photography:

(*3*) Relative orientation of the *zero-base model* (see Fig.8) requires the use of an *identical ray equation* of the form:

$$\lambda \begin{pmatrix} u' \\ v' \\ w' \end{pmatrix} = \begin{pmatrix} u'' \\ v'' \\ w'' \end{pmatrix}$$

OFEK07668

The factor λ allows for a variation in photoscale between both photographs involved.

The introduction of this condition presupposes that the two convergent cameras have a negligible separation of their nodal points. If this is not the case (or if they are not properly synchronized) the equation deviates to:

(*a*) Another coplanarity equation:

$$\begin{vmatrix} cx & cy & cz \\ u' & v' & w' \\ u'' & v'' & w'' \end{vmatrix} = 0$$

it gives a solution close to indeterminate due to the small base c involved.

(*b*) Therefore it is more appropriate to use an approximately known vector c to correct the equation:

$$\lambda \begin{pmatrix} u' \\ v' \\ w' \end{pmatrix} = \begin{pmatrix} u'' \\ v'' \\ w'' \end{pmatrix} + \begin{pmatrix} cu \\ cv \\ cw \end{pmatrix} \cdot \frac{f}{z}$$

where cu, cv and cw are functions of the orientation of the camera, but due to the small magnitude of c they can be approximated as:

$$\begin{pmatrix} cu \\ cv \\ cw \end{pmatrix} = \begin{pmatrix} \bar{a}_1 & \bar{a}_2 & \bar{a}_3 \\ \bar{b}_1 & \bar{b}_2 & \bar{b}_3 \\ \bar{c}_1 & \bar{c}_2 & \bar{c}_3 \end{pmatrix} \begin{pmatrix} cx \\ cy \\ cz \end{pmatrix} \approx \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} cx \\ cy \\ cz \end{pmatrix} \approx \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

in which $\bar{a}_1$ to $\bar{c}_3$ are coefficients of the rotational matrix of camera 1 in which φ has been replaced by $(\varphi' - \varphi_0')$; cx, cy and cz are the components of the nodal point separation, c_0 in case of an orientation $\varphi' - \varphi_0' = \omega' = k' = 0$ of camera 1.

(*c*) In the zero-base model two independent angular conditions can also be enforced assuming that the camera calibration of the cameras relative to each other is known a priori and that it remains stable during the operation. A constant convergence angle is enforced by the equation $c' \times c'' = \cos \gamma$.

One of the following conditions can also be enforced:

$$a' \times a'' = \cos \alpha$$

$$b' \times b'' = \cos \beta$$

0800

OFEK07669

in which α and β are space angles between the respective photocoordinate axes unit vectors a', a'', b', b'' and c', c''. They can also be included as additional unknowns. This has the following advantage: when including them as unknown parameters it becomes possible to weight them as hybrid measurements from $0 - \infty$ (SCHMID and SCHMID, 1965), if the adjustment is set up in the form of condition equations treating every parameter as an observation.

A point common to two convergent models results in two coplanarity conditions, two intersecting ray conditions, one identical ray equation in corrected or uncorrected form, one convergence angle condition and one other angular condition for the relative orientation of the twin cameras. The last two conditions are only justified if the camera mount can be considered sufficiently stable not to be affected by vibrating deformations which would cause varying deflections of the exposure direction during the flight.

Linearization of conditions

The linearized equations valid for a point common to two convergent models are summarized as follows:

(*1*) Coplanarity equation for the left model, in which the left exposure station is fixed as well as the base component bx' between left- and right-center exposure station:

$$\begin{vmatrix} bx' & by_0' & bz_0' \\ u' & v' & w' \\ 0 & -w_0'' & v_0'' \end{vmatrix} \alpha_1'' + \begin{vmatrix} bx' & by_0' & bz_0' \\ u' & v' & w' \\ w_0'' & 0 & u_0'' \end{vmatrix} \alpha_2'' + \begin{vmatrix} bx' & by_0' & bz_0' \\ u' & v' & w' \\ -v_0'' & u_0'' & 0 \end{vmatrix} \alpha_3''$$

$$+ \begin{vmatrix} w' & u' \\ w_0'' & u_0'' \end{vmatrix} dby' + \begin{vmatrix} u' & v' \\ u_0'' & v_0'' \end{vmatrix} dbz' +$$

$$+ \begin{vmatrix} w' & u' \\ w_0'' & u_0'' \end{vmatrix} by_0' + \begin{vmatrix} u' & v' \\ u_0'' & v_0'' \end{vmatrix} bz_0' + \begin{vmatrix} v' & w' \\ v_0'' & w_0'' \end{vmatrix} bx' =$$

$$- \begin{vmatrix} bx' & by_0' & bz_0' \\ a_1' & a_2' & a_3' \\ u_0'' & v_0'' & w_0'' \end{vmatrix} (dx' + v_x') - \begin{vmatrix} bx' & by_0' & bz_0' \\ b_1' & b_2' & b_3' \\ u_0'' & v_0'' & w_0'' \end{vmatrix} (dy' + v_y')$$

OFEK07670

$$-\begin{vmatrix} bx' & by_0' & bz_0' \\ u' & v' & w' \\ a_{10}'' & a_{20}'' & a_{30}'' \end{vmatrix}(dx'' + v_x'') - \begin{vmatrix} bx' & by_0' & bz_0' \\ u' & v' & w' \\ b_1'' & b_2'' & b_3'' \end{vmatrix}(dy'' + v_y'')$$

Allowance is made for eventual correction of known interior-orientation errors dx', dy', dx'' and dy''. The terms v_x', v_y', v_x'' and v_y'' are the errors in plate-coordinate measurement of the left- and the right-center photograph.

The linearized coplanarity condition for the second model becomes:

$$\begin{vmatrix} bx_0'' & by_0'' & bz_0'' \\ 0 & -w_0''' & v_0''' \\ u_0'''' & v_0'''' & w_0'''' \end{vmatrix} a_1'''' + \begin{vmatrix} bx_0'' & by_0'' & bz_0'' \\ u_0''' & v_0''' & w_0''' \\ 0 & -w_0'''' & v_0'''' \end{vmatrix} a_1'''$$

$$+\begin{vmatrix} bx_0'' & by_0'' & bz_0'' \\ w_0''' & 0 & -u_0''' \\ u_0'''' & v_0'''' & w_0'''' \end{vmatrix} a_2'''' + \begin{vmatrix} bx_0'' & by_0'' & bz_0'' \\ u_0''' & v_0''' & w_0''' \\ w_0'''' & 0 & -u_0'''' \end{vmatrix} a_2'''$$

$$+\begin{vmatrix} bx_0'' & by_0'' & bz_0'' \\ -v_0''' & u_0''' & 0_0''' \\ u_0'''' & v_0'''' & w_0'''' \end{vmatrix} a_3'''' + \begin{vmatrix} bx_0'' & by_0'' & bz_0'' \\ u_0''' & v_0''' & w_0''' \\ -v_0'''' & u_0'''' & 0'''' \end{vmatrix} a_3'''$$

$$+\begin{vmatrix} w_0''' & u_0''' \\ w_0'''' & u_0'''' \end{vmatrix} dby'' + \begin{vmatrix} u_0''' & v_0''' \\ u_0'''' & v_0'''' \end{vmatrix} dbz'' + \begin{vmatrix} v_0''' & w_0''' \\ v_0'''' & w_0'''' \end{vmatrix} dbx''$$

$$-\begin{vmatrix} w_0''' & u_0''' \\ w_0'''' & u_0'''' \end{vmatrix} dby'' - \begin{vmatrix} u_0''' & v_0''' \\ u_0'''' & v_0'''' \end{vmatrix} dbz'$$

$$+\begin{vmatrix} bx_0'' & by_0'' & bz_0'' \\ u_0''' & v_0''' & w_0''' \\ u_0'''' & v_0'''' & w_0'''' \end{vmatrix} =$$

OFEK07671

$$-\begin{vmatrix} bx_0'' & by_0'' & bz_0'' \\ a_{1_0}''' & a_{2_0}''' & a_{3_0}''' \\ u_0'''' & v_0'''' & w_0'''' \end{vmatrix}(dx'''' + v_x'''') - \begin{vmatrix} bx_0'' & by_0'' & bz_0'' \\ b_{1_0}''' & b_{2_0}''' & b_{3_0}''' \\ u_0'''' & v_0'''' & w_0'''' \end{vmatrix}(dy'''' + v_y'''')$$

$$-\begin{vmatrix} bx_0'' & by_0'' & bz_0'' \\ u_0''' & v_0''' & w_0''' \\ a_{1_0}'''' & a_{2_0}'''' & a_{3_0}'''' \end{vmatrix}(dx'''' + v_x'''') - \begin{vmatrix} bx_0'' & by_0'' & bz_0'' \\ u_0''' & v_0''' & w_0''' \\ b_{1_0}'''' & b_{2_0}'''' & b_{3_0}'''' \end{vmatrix}(dy'''' + v_y'''')$$

in which:

$$bx_0'' = x_0'''' - x_0''' = bx;$$

$$by_0'' = y_0'''' - y_0''' = dby_0'';$$

$$bz_0'' = z_0'''' - z_0''' = dbz_0''.$$

The terms with three primes refer to the left-center camera (model 2) and to the base components, and the terms with four primes to the right camera (of the second model).

(2) *Intersection of three rays.* The linearization of this equation starts from:

$$\begin{vmatrix} bx' & (bz_0' + dbz') \\ u' & w' \end{vmatrix}\begin{vmatrix} (u_0'' + du'')(w_0'' + dw'') \\ (u_0''' + du''')(w_0''' + dw''') \end{vmatrix}$$

$$+\begin{vmatrix} (bx'' + dbx'')(bz_0'' + dbz'') \\ (u_0''' + du''')(w_0''' + dw''') \end{vmatrix}\begin{vmatrix} (u_0'' + du'')(w_0'' + dw'') \\ u' \qquad\qquad w' \end{vmatrix} = 0$$

Setting:

$$\bar{a} = -\begin{vmatrix} u_0'' & w_0'' \\ u_0''' & w_0''' \end{vmatrix} u' \qquad \bar{\theta} = \begin{vmatrix} u_0'' & w_0'' \\ u_0''' & w_0''' \end{vmatrix}$$

$$\bar{\beta} = \begin{vmatrix} u_0'' & w_0'' \\ u' & w' \end{vmatrix} w_0'' \qquad \bar{k} = \begin{vmatrix} u_0'' & w' \\ u' & w_0'' \end{vmatrix}$$

$$\bar{\gamma} = -\begin{vmatrix} u_0'' & w_0'' \\ u' & w' \end{vmatrix} u_0''' \qquad \bar{\lambda} = \begin{vmatrix} bx_0' & bz_0' \\ u_0'' & w_0'' \end{vmatrix}$$

OFEK07672

$$\bar{\delta} = \left(\begin{vmatrix} bx' & bz_0' \\ u' & w' \end{vmatrix} w_0''' + \begin{vmatrix} bx_0'' & bz_0'' \\ u_0''' & w_0''' \end{vmatrix} w' \right) \qquad \bar{\mu} = \begin{vmatrix} bx' & bz_0' \\ u' & w' \end{vmatrix}$$

$$\bar{\varepsilon} = -\left(\begin{vmatrix} bx' & bz_0' \\ u' & w' \end{vmatrix} u_0''' + \begin{vmatrix} bx_0'' & bz_0'' \\ u_0''' & w_0''' \end{vmatrix} u' \right)$$

$$\bar{\zeta} = -\left(\begin{vmatrix} bx' & bz_0' \\ u' & w' \end{vmatrix} w_0'' + \begin{vmatrix} u_0'' & w_0'' \\ u' & w' \end{vmatrix} bz_0'' \right)$$

$$\bar{\eta} = +\left(\begin{vmatrix} bx' & bz_0' \\ u' & w' \end{vmatrix} u_0'' + \begin{vmatrix} u_0'' & w_0'' \\ u' & w' \end{vmatrix} bx_0'' \right)$$

the linearized equation reads:

$$\begin{aligned}
&\bar{\varepsilon}\, v_0''\, \alpha_1'' + (\bar{\delta}\, w_0'' - \bar{\varepsilon}\, w_0'')\, \alpha_2'' - \bar{\delta}\, v_0''\, \alpha_3'' \\
&+ \bar{\eta}\, v_0'''\, \alpha_1''' + (\bar{\zeta}\, w_0''' - \bar{\eta}\, u_0''')\, \alpha_2''' - \bar{\zeta}\, v_0'''\, \alpha_3''' \\
&+ \bar{\alpha}\, dbz' + \bar{\beta}\, dbx'' - \bar{\gamma}\, dbz'' = \\
&= -\left(\begin{vmatrix} bx' & bz_0' \\ a_{10}' & a_{30}' \end{vmatrix} \bar{\Theta} + \bar{\lambda} \begin{vmatrix} u_0'' & w_0'' \\ a_{10}' & a_{30}' \end{vmatrix} \right) (dx' + v_x') \\
&- \left(\begin{vmatrix} bx' & bz_0' \\ b_{10}' & b_{30}' \end{vmatrix} \bar{\Theta} + \bar{\lambda} \begin{vmatrix} u_0'' & w_0'' \\ b_{10}' & b_{30}' \end{vmatrix} \right) (dy' + v_y') \\
&- \left(\begin{vmatrix} a_{10}'' & a_{30}'' \\ u_0''' & w_0''' \end{vmatrix} \bar{\mu} + \bar{\lambda} \begin{vmatrix} a_{10}'' & a_{30}'' \\ u' & w' \end{vmatrix} \right) (dx'' + v_x'') \\
&- \left(\begin{vmatrix} b_{10}'' & b_{30}'' \\ u_0''' & w_0''' \end{vmatrix} \bar{\mu} + \bar{\lambda} \begin{vmatrix} b_{10}'' & b_{30}'' \\ u' & w' \end{vmatrix} \right) (dy'' + v_y'') \\
&- \left(\begin{vmatrix} u_0'' & w_0'' \\ a_{10}''' & a_{30}''' \end{vmatrix} \bar{\mu} + \bar{k} \begin{vmatrix} bx_0'' & bz_0'' \\ a_{10}''' & a_{30}''' \end{vmatrix} \right) (dx''' + v_x''') \\
&- \left(\begin{vmatrix} u_0'' & w_0'' \\ b_{10}''' & b_{30}''' \end{vmatrix} \bar{\mu} + \bar{k} \begin{vmatrix} bx_0'' & bz_0'' \\ b_{10}''' & b_{30}''' \end{vmatrix} \right) (dy''' + v_y''')
\end{aligned}$$

0804 OFEK07673

The terms with one prime refer to the first camera of the first model, the terms with two and three prime can be varied since two such equations exist for one point.

(3) *Identical ray equation.* It is linearized from:

$$\lambda \begin{pmatrix} u' \\ v' \\ w' \end{pmatrix} = \begin{pmatrix} u_0'' \\ v_0'' \\ w_0'' \end{pmatrix} + \begin{pmatrix} du'' \\ dv'' \\ dw'' \end{pmatrix}$$

$$\begin{pmatrix} du'' \\ dv'' \\ dw'' \end{pmatrix} = \begin{pmatrix} \frac{\partial u''}{\partial \alpha_1} & \frac{\partial u''}{\partial \alpha_2} & \frac{\partial u''}{\partial \alpha_3} \\ \frac{\partial v''}{\partial \alpha_1} & \frac{\partial v''}{\partial \alpha_2} & \frac{\partial v''}{\partial \alpha_3} \\ \frac{\partial w''}{\partial \alpha_1} & \frac{\partial w''}{\partial \alpha_2} & \frac{\partial w''}{\partial \alpha_3} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$

if camera one-prime is considered oriented:

$$\lambda u' = u_0'' + w_0'' \alpha_2 - v_0'' \alpha_3$$

$$\lambda v' = v_0'' - w_0'' \alpha_1 + u_0'' \alpha_3$$

The third equation:

$$\lambda w' = w_0'' + v_0'' \alpha_1 - u_0'' \alpha_2$$

is not independent. The terms with a prime are associated with the second camera of the first model and the terms with two prime with the first camera of the second model. The linearized equations become for the case that also the camera one prime needs to be adjusted:

$$u_0' + a_1' (dx' + v_x') + b_1' (dy' + v_y') - w_0' \alpha_2' - v_0' \alpha_3' =$$
$$u_0'' + w_0'' \alpha_2'' - v_0'' \alpha_3'' + a_1'' (dx'' + v_x'') + b_1'' (dy'' + v_y'') + (cu) \frac{f}{z}$$

$$v_0' + a_2' (dx' + v_x') + b_2' (dy' + v_y') - w_0' \alpha_1' + u_0' \alpha_3' =$$
$$v_0'' - w_0'' \alpha_1'' \alpha_3'' + a_2'' (dx'' + v_x'') + b_2'' (dy'' + v_y'') + (cv) \frac{f}{z}$$

For the solution of the zero-base model $cu \frac{f}{z}$, $cv \frac{f}{z}$, dx', dy', dx'', dy'' are

OFEK07674

considered known. One then limits oneself to the determination of four unknowns which is most conveniently made by measurements at three points (six equations) along a line in the zero-base model.

(4) *Fixed angle conditions.* The equation for fixed angles may be linearized as:

$$(c_0' + dc')(c_0'' + dc'') - \cos\gamma = 0$$

$$dc'' = \begin{pmatrix} \frac{\partial c_1''}{\partial \omega_1''} & \frac{\partial c_1''}{\partial \omega_2''} & \frac{\partial c_1''}{\partial \omega_3''} \\ \frac{\partial c_2''}{\partial \omega_1''} & \frac{\partial c_2''}{\partial \omega_2''} & \frac{\partial c_2''}{\partial \omega_3''} \\ \frac{\partial c_3''}{\partial \omega_1''} & \frac{\partial c_3''}{\partial \omega_2''} & \frac{\partial c_3''}{\partial \omega_3''} \end{pmatrix} \begin{pmatrix} \omega_1'' \\ \omega_2'' \\ \omega_3'' \end{pmatrix} = \begin{pmatrix} 0 & -c_3'' & c_2'' \\ c_3'' & 0 & -c_1'' \\ -c_2'' & c_1 & 0 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix}$$

and therefore:

$$\begin{aligned} &c_1' c_{01}'' + c_2' c_{02}'' - \cos\gamma + (c_2' c_3'' - c_3' c_2'')\,\omega_1' + (c_3' c_1'' - c_1' c_3'')\,\omega_2' \\ &+ (c_1' c_2'' - c_2' c_1'')\,\omega_3' + (c_3' c_2'' - c_2' c_3'')\,\omega_1'' + (c_3' c_1'' - c_1' c_3'')\,\omega_2'' \\ &+ (c_1' c_2'' - c_2' c_1'')\,\omega_3'' = 0 \end{aligned}$$

Inclusion of this type of an equation necessitates the adjustment of the system as observation equations with conditions, since no observations are contained in them. Equations for angle α or β are analogous.

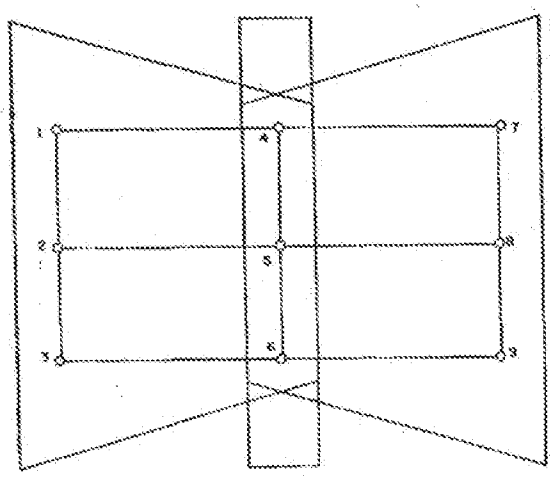


Fig.7. Minimum point distribution for the analytical orientation of two convergent models.

0806 OFEK07675

Adjustment of the "quadruplet" and strip formation

For the standard-measurement case in triangulation the conditions are applied to usually a minimum of nine points in characteristic locations (Fig.7) for which the following considerations are valid. An increased number of points at such characteristic locations only increases the number of equations but not the parameters.

There will be six coplanarity equations of the first type to points 1–6 and six coplanarity conditions of the second type to points 4–9. There will exist six intersection conditions, two each for points 4, 5 and 6. There are six identical ray equations, two each for points 4, 5 and 6; and in addition two fixed angle conditions can be introduced giving a total of 26 equations. These will be used for the adjustment of the fourteen parameters; the nine rotations of three cameras and the five base components bx', bz', bx'', by'' and bz''.

Since the rotations α_1, α_2, α_3 of each camera are around rotated axes the α_i have only meaning when used to improve the introduced approximation $\mathbf{A}_0$ of the final matrix $\mathbf{A}$ for each camera:

$$\mathbf{A} = d\mathbf{A} \cdot \mathbf{A}_0$$

or:

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = \begin{pmatrix} 1 & -\alpha_3 & +\alpha_2 \\ +\alpha_3 & 1 & -\alpha_1 \\ -\alpha_2 & +\alpha_1 & 1 \end{pmatrix} \begin{pmatrix} a_{01} & b_{01} & c_{01} \\ a_{02} & b_{02} & c_{02} \\ a_{03} & b_{03} & c_{03} \end{pmatrix}$$

the α_i are used in the calculation of $d\mathbf{A}$. $d\mathbf{A}$ results in $\mathbf{A}$ with which:

$$\mathbf{U} = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$$

is calculated. This procedure is iterated. Rotations φ, ω, $\varkappa$ of any sequence may however be deduced from the elements of matrix $\mathbf{A}$.

There are altogether 48 photocoordinate observations in the standard measurement case.

Fig.8 illustrates the properties of the combined matrix which has a general form of:

0807

OFEK07676

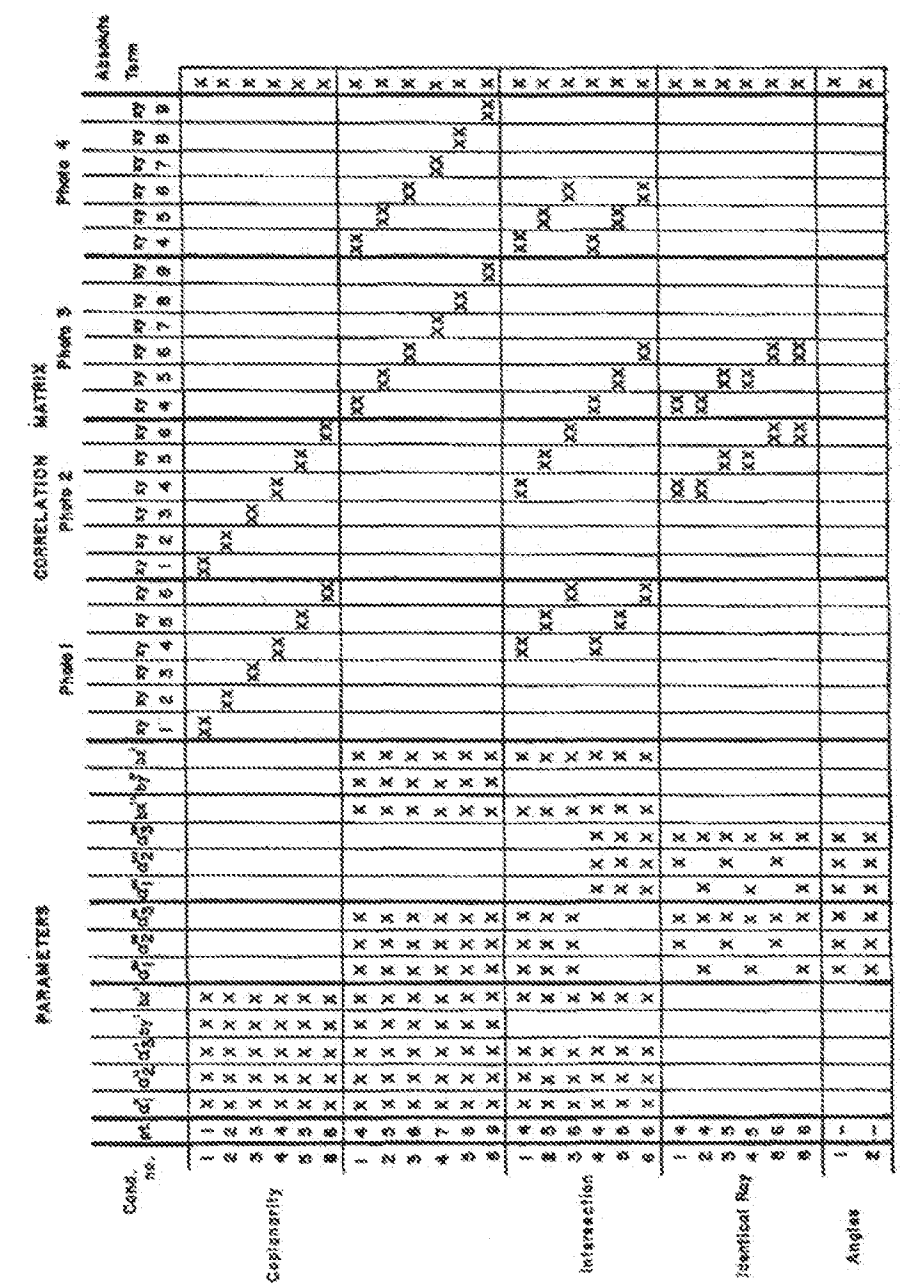


Fig. 8. Schematic diagram of equations and terms of a "quadruplet".

OFEK07677

$$\underset{(24,14)}{\mathbf{A}}\ \underset{(14,1)}{\mathbf{X}} - \underset{(24,1)}{\mathbf{L}} = \underset{(24,48)}{\mathbf{B}}\ \underset{(48,1)}{\mathbf{V}}$$

$$\underset{(2,14)}{\mathbf{C}}\ \underset{(14,1)}{\mathbf{X}} - \underset{(2,1)}{\mathbf{W}} = 0$$

$\mathbf{X}$ represents the parameter vector, $\mathbf{A}$ the parameter coefficient matrix, $\mathbf{L}$ the absolute term, $\mathbf{V}$ the correction vector to the observations and $\mathbf{B}$ the correlation matrix. If $\mathbf{P}$ is another correlation matrix expressing the correlation between measured x' and y' photocoordinates, a diagonal weight matrix or a unit matrix, then the adjustment problem is given by the function F, which has to be minimized:

$$F = \mathbf{V}^T\mathbf{P}\mathbf{V} - 2\mathbf{K}^T(\mathbf{B}\mathbf{V} - \mathbf{A}\mathbf{X} + \mathbf{L}) + 2\mathbf{K}_2^T(\mathbf{C}\mathbf{X} - \mathbf{W})$$

$$\frac{\partial F}{\partial \mathbf{V}} = \min = 2\mathbf{V}^T\mathbf{P} - 2\mathbf{K}_1^T\mathbf{B} = 0; \text{ or } \mathbf{V} = \mathbf{P}^{-1}\mathbf{B}^T\mathbf{K}_1$$

$$\frac{\partial F}{\partial \mathbf{X}} = \min = +2\mathbf{K}_1^T\mathbf{A} + 2\mathbf{K}_2^T\mathbf{C} = 0; \text{ or } \mathbf{A}^T\mathbf{K}_1 = -\mathbf{C}^T\mathbf{K}_2$$

$\mathbf{V}$ with the matrix gives:

$$\mathbf{A}\mathbf{X} - \mathbf{L} = \mathbf{B}\mathbf{P}^{-1}\mathbf{B}\mathbf{K}_1$$

this results in:

$$\mathbf{K}_1 = (\mathbf{B}\mathbf{P}^{-1}\mathbf{B}^T)^{-1}(\mathbf{A}\mathbf{X} - \mathbf{L})$$

If no matrix $\mathbf{C}$ with angular conditions is introduced, $\mathbf{K}_2 = 0$ and $\mathbf{A}^T\mathbf{K}_1 = 0$. Therefore:

$$\mathbf{X}_0 = [\mathbf{A}^T(\mathbf{B}\mathbf{P}^{-1}\mathbf{B}^T)^{-1}\mathbf{A}]^{-1}[\mathbf{A}^T(\mathbf{B}\mathbf{P}^{-1}\mathbf{B}^T)^{-1}\mathbf{L}]$$

If $\mathbf{C}$ is introduced:

$$\mathbf{X} = [\mathbf{A}^T(\mathbf{B}\mathbf{P}^{-1}\mathbf{B}^T)^{-1}\mathbf{A}]^{-1}[\{\mathbf{A}^T(\mathbf{B}\mathbf{P}^{-1}\mathbf{B}^T)^{-1}\mathbf{L}\} - \mathbf{C}^T\mathbf{K}_2]$$

SCHMID (1965) gives a solution in which the final $\mathbf{X}$, under consideration of $\mathbf{C}$, may be obtained as:

$$\mathbf{X} = \mathbf{X}_0 + \delta\mathbf{X}$$

$$\delta\mathbf{X} = -[\mathbf{A}^T(\mathbf{B}\mathbf{P}^{-1}\mathbf{B}^T)^{-1}\mathbf{A}]^{-1}\mathbf{C}^T\mathbf{K}_2$$

OFEK07678

or:

$$K_s = -C^T [A^T (BP^{-1} B^T)^{-1} A] \delta X; \text{ and out of } C (X_0 + \delta X) - W = 0$$

follows:

$$\delta X = -C^{-1} (CX_0 - W)$$

$$\delta X = -[A^T (BP^{-1} B^T)^{-1} A]^{-1} C^T [C [A^T (BP^{-1} B^T)^{-1} A] C^T]^{-1} (CX_0 - W)$$

The main computational effort are the two matrix inversions of the correlation matrix (24 × 24), and of the parameter coefficients (14 × 14). The correlation matrix can be reduced further in size, but in any case a matrix inversion of a 24 × 24 system will not exceed the capacity of a small computer. It will even be sufficient to use an approximate correlation matrix for the general case of convergent photography with $\varphi = \gamma/2$ and $\omega = \varkappa = 0$ for all cases. The inverse can thus be precalculated for a standard arrangement of points. Once the vector **X** has been calculated rotational matrices **dA** and **A** can be formed and model coordinates **U** determined. After sufficient convergence of the solution is reached after a few iterations, the next "quadruplet" can be calculated *keeping the orientation of the second exposure station of the previous quadruplet fixed* (consisting of two cameras) and allowing for a readjustment of exposure station three. In this way the strip may be triangulated in a manner similar to Elassal's University of Illinois program operating with triplets of vertical photographs (ELASSAL, 1963; KARARA et al., 1964).

It is hoped that the error propagation may be significantly reduced in this way. The assembled strip may then be adjusted by conventional methods. Larger computers could use the output as good approximation for a rigorous block adjustment utilizing simultaneously all conditions according to SCHMID (1965). It is regretted that results of a practical test of the method presented cannot be given at this time.

Study of error analysis

For the study of error analysis it is easily possible to evaluate the well-known model orientation formulae for convergent photography as given by HALLERT (1954), ACKERMANN (1956) or KONECNY (1965) in conjunction with a correlation matrix between photocoordinates and model coordinates. This has the advantage that points in predetermined model locations can be easier introduced. It would also be possible to formulate a method of computation in this manner which would conform to a rigorous least squares adjustment.

0810 OFEK07679

REFERENCES

ACKERMANN, F., 1956. Zur rechnerischen Orientierung von Konvergentaufnahmen. *Bildmessung Luftbildwesen, Sonderh.*, 24: 24–37.

ARTHUR, D. W. G., 1956. Relative orientation of photographs taken from the same station. *Photogrammetric Eng.*, 22: 935–938.

BRANDENBERGER, A., 1957. Some considerations about error propagation in strip triangulations: attainable accuracy. *Photogrammetria:* 12: 78–85.

ELASSAL, A., 1963. Analytical aerotriangulation at the University of Illinois. *Photogrammetric Eng.*, 30: 199–206.

GRIFFIN, E. P., 1959. Vertical photography for aerotriangulation. *Photogrammetric Eng.*, 25.

HALLERT, B., 1954. Some remarks concerning the theory of errors for convergent aerial pictures in comparison with near vertical pictures. *Photogrammetric Eng.*, 20: 749–757.

HARRIS, W. D., TEWINKEL, G. C. and WHITTEN, C. A., 1962. Analytic aerotriangulation. *U.S. Coast Geodetic Surv., Tech. Bull.*, 21.

KARARA, H. M. and ELASSAL, A. A., 1964. Simultaneous relative orientation of multiple cameras. *J. Am. Soc. Civil Engrs.*, 90: 109–116.

KONECNY, G., 1962. Aerotriangulation mit Konvergentaufnahmen. *Veröffentl. Deut. Geodätischen Komm.*, C (47): 50 S.

KONECNY, G., 1965. Interior orientation and convergent photography. *Photogrammetric Eng.*, 31: 625–634.

LEHMANN, G., 1963. Die Verwendung von Bildpaaren und Bildtripeln bei der Aerotriangulation von Bildstreifen. *Z. Vermessungswesen*, 92: 485–492.

PENNINGTON, J. T., 1954. Aerotriangulation with convergent photography. *Photogrammetric Eng.*, 20: 76–81.

RINNER, K., 1957. Die analytische Lösung des Koppelbildanschlusses. *Bildmessung Luftbildwesen*, 25: 46–49.

ROSENFIELD, G. H., 1957. The precise determination of relative orientation between two photographs taken at a single exposure station, and a measure of its accuracy. *Photogrammetric Eng.*, 27: 948–952.

SCHMID, H. H., 1956. An analytical treatment of the problem of triangulation by stereophotogrammetry, *Photogrammetria*, 12: 67–77; 91–116.

SCHMID, H. H. and SCHMID E., 1965. A generalized least squares solution for hybrid measuring systems. *Can. Surveyor*, 19: 27–41.

SCHUT, G. H., 1964. The N.R.C. program for analytical aerial triangulation on the I.B.M. 1620 and I.B.M. 650. *Natl. Res. Council, Publ.* 8310.

THEIS, J. B., 1958. Increased base: height ratio. *Photogrammetric Eng.*, 24: 127–132.

THOMPSON, E. H., 1957. Convergent obliques in analytical aerial triangulation. *Photogrammetric Record*, 1: 219–223.

VON GRUBER, O., 1930. *Ferienkurs für Photogrammetrie*. Wittwer Verlag, Stuttgart. 468 S.

OFEK07680

Reprinted from
PHOTOGRAMMETRIC ENGINEERING
JULY
1965

PROF. G. KONECNY*
Univ. of New Brunswick
Fredericton, N. B., Canada

Interior Orientation and Convergent Photography

ABSTRACT: *Convergent photography, though awkward in use, yields a significant theoretical accuracy increase over vertical photography. The practical accuracy is lower than theoretically expected, mainly due to errors of interior orientation. These affect the accuracy of relative orientation for convergent photographs. More responsive procedures are suggested for orientation and bridging. Practical test results indicate that accuracy can thus be considerably improved.*

INTRODUCTION

The idea of convergent photography is as old as stereo-photogrammetry. Convergent photographs were used from balloons around 1900, (S. Finsterwalder, 1903, pp. 225–260), a long time before the airplane would permit a mechanized sequence of vertical photographs.

In 1926, a convergent photography procedure by twin cameras was developed to increase the intersection accuracy of corresponding photogrammetric rays. By the introduction of wide-angle photography in 1934, however, an increase of accuracy could also be achieved, without having to give up the simplicity of viewing and measuring from vertical photographs.

Despite the difficulties in stereovision, due to wide variations of scale and despite the complexity of the system, wide-angle convergent photography was reintroduced in the 1950's by the U. S. Geological Survey to further economize plotting and bridging. Super-wide angle development again stopped the widespread use of convergent photography.

Today not primarily economic considerations but increased accuracy still open perspectives for convergent photography.

It has been shown that for special projects such as flood control surveys and for certain engineering projects, extreme accuracies, unachieved by vertical photography, could be attained, particularly from low altitudes.

The peculiarities and the awkwardness of working with convergent photography naturally have led only to limited acceptance of convergent procedures. On the other hand, though, theoretical considerations have always shown (Hallert, 1954) that convergent photography should yield a sizeable increase in accuracy, this has not been found in practice as Table 1 shows.



PROF. G. KONECNY

* Paper presented at the Semiannual Meeting of the American Society of Photogrammetry, 1963.

0812

OFEK07681

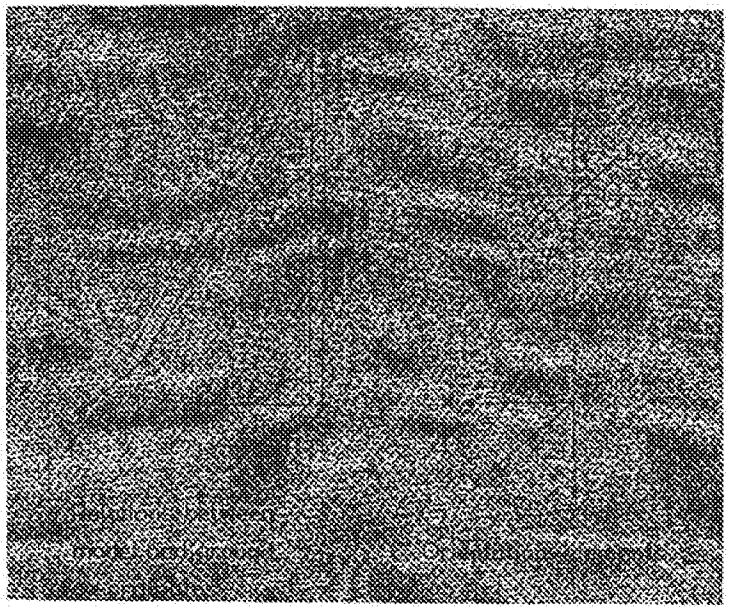


FIG. 1. Relations between Model and Ground Coordinates and Orientation Elements.

Tests conducted by AMS (Theis 1958) and by ERDL (Pennington 1954) show for the most part only an insignificant accuracy increase for convergent photography, or a decrease of accuracy for convergent bridging procedures.

TABLE 1

PRACTICAL ACCURACY REACHED WITH CONVERGENT PHOTOGRAPHY

Author	*Year*	*Instrument*	*Procedure*	m_o in °/₀₀ of h	m_h in °/₀₀ of h
Pennington ERDL	1954	Multiplex	Bridging 6 models	±1.67 °/₀₀ h	±0.71 °/₀₀ h
Theis AMS	1958	Kelsh	WA[1] vert. points		±0.17
			WA conv. points		±0.16
		Balplex	WA vert. points		±0.24
			WA conv. points		±0.21
		Kelsh	WA vert. contours		±0.24
			WA conv. contours		±0.18
		Balplex	WA vert. contours		±0.27
			WA conv. contours		±0.22
		C8	WA vert. bridging 15 models		±0.18
		C8	WA conv. bridging 8 models		±0.33
Griffin ERDL	1959	C8 and BL 720	WA vert. bridging 16 models	±1.55	±0.69
		C8 and BL 720	WA conv. bridging 8 models	±1.59	±0.54

[1] WA = wide-angle.

OFEK07682

More recent investigations have led to the cause of this (E. P. Griffin, 1959) (G. Konecny, 1962), but the right conclusions have generally not yet been drawn. Particularly the errors in interior orientation have led to difficulties in attaining the expected theoretical accuracy.

Interior Orientation Errors and Model Deformations

If x, y, h designate ground co-ordinates, x', y', f photo co-ordinates, x_R, y_R, f_R, x_R', y_R', z_R' rotated ground and photo co-ordinates, and A an orthogonal rotational matrix, we have the well-known perspective relations:

$$\begin{pmatrix} x \\ y \\ h \end{pmatrix} = A \begin{pmatrix} x_R \\ y_R \\ z_R \end{pmatrix}; \qquad \begin{pmatrix} x_R' \\ y_R' \\ z_R' \end{pmatrix} = A \begin{pmatrix} x' \\ y' \\ f \end{pmatrix};$$

$$x = \frac{h}{z_R} x_R; \qquad y = \frac{h}{z_R} y_R; \qquad x' = \frac{f}{z_R'} x_R'; \qquad y' = \frac{f}{z_R'} y_R'$$

from which we obtain by linearization and differentiation the complete differential relations between model co-ordinates (x, y, h) and exterior ($d\phi$, $d\omega$, $d\kappa$) as well as interior orientation elements (dx', dy', df). (See Figure 1.)

For a rotational sequence ϕ, ω, κ and the signs of the C8 we obtain:

$$dx = h\left(1 + \frac{x^2}{h^2}\right) d\phi + \frac{y}{h}(h \sin\phi - x \cos\phi) d\omega - \frac{y}{h}(h \cos\phi + x \sin\phi) d\kappa$$
$$- \frac{x}{h} dbz + dbx + \frac{1}{hf}(h \cos\phi + x \sin\phi)^2 dx'$$
$$+ \frac{1}{hf}(h \sin\phi - x \cos\phi)(h \cos\phi + x \sin\phi) df;$$

$$dy = \frac{xy}{h} d\phi - \left[x \sin\phi + h \cos\phi\left(1 + \frac{y^2}{h^2}\right)\right] d\omega + \left[x \cos\phi - h \sin\phi\left(1 + \frac{y}{h^2}\right)\right] d\kappa$$
$$- \frac{y}{h} dbz - dby + \frac{1}{hf}\left[y \sin\phi(x \sin\phi + h \cos\phi)\right] dx'$$
$$+ \frac{1}{f}(x \sin\phi + h \cos\phi) dy' - \frac{1}{hf} y \cos\phi(x \sin\phi + h \cos\phi) df$$

for use with "base outside" we get:

$$py = dy_R - dy_L$$
$$px = dx_R - dx_L; \qquad dh = \frac{h}{b} px,$$

characterizing model deformations. A graphical representation shows that elements $d\omega$, $d\kappa$, dx', dy', df act differently in convergent and near vertical photographs, while $d\phi$ acts the same. (See Figure 2.)

As is well known, interior orientation errors are not harmful to vertical photography except for translations of the model in the x, y and z directions. For convergent photographs, however, an erroneous relative orientation will result, since the y parallaxes introduced by dx', dy' and df will cause an erroneous compensation of these parallaxes by other elements:

OEEK07683

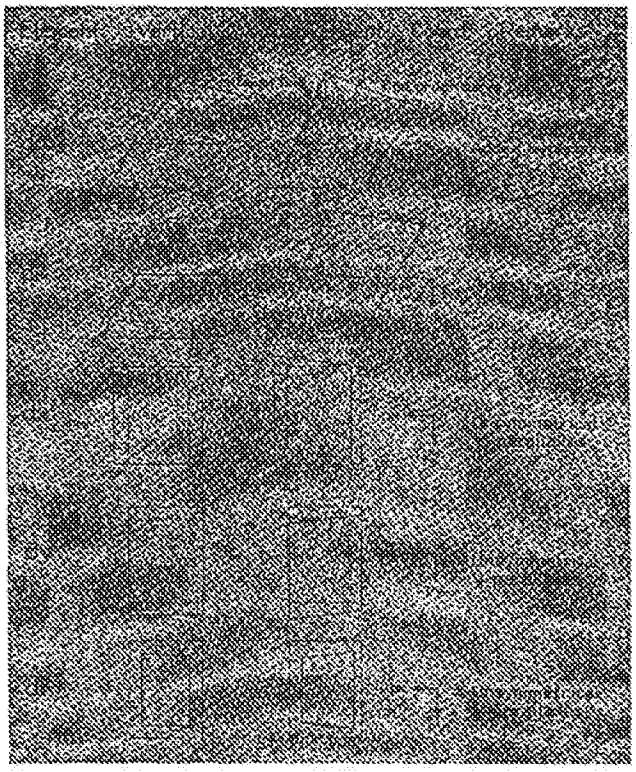


FIG. 2. Model Deformation for Vertical and Convergent Photography.

0.1 mm dx' will be compensated by $+0.15^s\ d\phi - 0.05$ mm db_x

0.1 mm dy' will be compensated by $-0.67^s\ d\kappa - 0.17^s\ d\omega$

0.1 mm df will be compensated by $-0.69^s\ d\phi - 0.22$ mm db_z.

Analytically this can be expressed as follows: Using the model co-ordinates of the orientation points 1 to 6 in the formula

$$
\begin{aligned}
py = {} & \frac{1}{f_R k}(xy\sin^2\phi + yk\sin\phi\cos\phi)dx_R' + \frac{1}{f_R k}(xk\sin\phi + k^2\cos\phi)dy_R' \\
& + \frac{1}{f_R k}(-xy\sin\phi\cos\phi - yk\cos^2\phi)df_R \\
& + \frac{1}{f_L k}(-(x-b)y\sin^2\phi - yk\sin\phi\cos\phi)dx_L' \\
& + \frac{1}{f_L k}(-(x-b)k\sin\phi - k^2\cos\phi)dy_L' \\
& + \frac{1}{f_L k}((x-b)y\sin\phi\cos\phi + yk\cos^2\phi)df_L
\end{aligned}
$$

OFEK07684

we obtain six expressions for parallaxes p_1 to p_6 which can be used to obtain the exterior orientation elements in terms of interior orientation errors:

$$d\phi_R = -\frac{h}{2bd}\left((p_3 - p_2) - (p_4 - p_6)\right)$$

$$= \sin^2\phi\left(\frac{dx_R'}{f_R} - \frac{dx_L'}{f_L}\right) - \sin\phi\cos\phi\left(\frac{df_R}{f_R} - \frac{df_L}{f_L}\right);$$

$$db z_R = +\frac{h}{2d}(p_4 - p_5) = h\sin\phi\cos\phi\left(\frac{dx_R'}{f_R}\right) - (b\sin^2\phi - h\sin\phi\cos\phi)\left(\frac{dx_L'}{f_L}\right)$$

$$+ h\cos^2\phi\left(\frac{df_R}{f_R}\right) + (b\sin\phi\cos\phi - h\cos^2\phi)\left(\frac{df_L}{f_L}\right);$$

$$db y_R = +\frac{b^2}{4d^2}\left[(p_3 + p_4 + p_5 + p_6 - 2p_1 - 2p_2)\right] + \frac{1}{6}\left[(p_1 - p_2) + (p_3 - p_4) + (p_5 - p_6)\right]$$

$$-\frac{1}{2}\left[(p_3 + p_5)\right] = h\cos\phi\left(\frac{dy_R'}{f_R} - \frac{dy_L'}{f_L}\right);$$

$$dx_R = -\frac{h}{4d^2}\sin\phi(p_3 + p_4 + p_5 + p_6 - 2p_1 - 2p_2)$$

$$+\frac{1}{3b}\cos\phi(p_1 - p_2 + p_3 - p_4 + p_5 - p_6)$$

$$= \sin\phi\cos\phi\left(\frac{dy_R'}{f_R} - \frac{dy_L'}{f_L}\right)$$

$$d\omega_R = -\frac{h}{4d^2}\cos\phi(p_3 + p_4 + p_5 + p_6 - 2p_1 - 2p_2)$$

$$-\frac{1}{3b}\sin\phi(p_1 - p_2 + p_3 - p_4 + p_5 - p_6)$$

$$= -\sin^2\phi\left(\frac{dy_R'}{f_R} - \frac{dy_L'}{f_L}\right)$$

The formulae show that only the differences $(dx_R' - dx_L')$, $(dy_R' - dy_L')$, $(df_R - df_L)$ act as error source. Therefore, systematic errors of interior orientation due to camera calibration have no influence for vertical photographs, since the same camera is used for all photos. Convergent photographs, however, use different cameras, and different films with different film shrinkage characteristics, which may cause severe strip bending in bridging but which may be compensated approximately by absolute orientation of the individual model.

From the above formulae weight coefficients are calculable, to study the effect of random errors of interior orientation. These weight coefficients are the elements of the inverse of the normal equation matrix. The normal equation matrix is formed by aid of the parallax equations at the chosen number of orientation points; for Hallert these are six. For example:

$$Q_{xx'} = \frac{2\sin^4\phi}{f^2}\frac{m_{x_L'}^2}{m_R^2} + \frac{2\sin^2\phi\cos^2\phi}{f^2}\frac{m_f^2}{m_c^2}$$

OFEK07685

TABLE 2

THE EFFECTS OF INTERIOR ORIENTATION ERRORS ON THE ACCURACY OF RELATIVE ORIENTATION

Column No.	1	2	3	4	5	6
	Convergent		Convergent		Vertical	
m_0	$\pm 6\ \mu$	$\pm 6\ \mu$	$\pm 20\ \mu$	$\pm 20\ \mu$	$\pm 20\ \mu$	$\pm 20\ \mu$
m_{dx}', m_{dy}'	$\pm 20\ \mu$	0	$\pm 100\ \mu$	0	$\pm 100\ \mu$	0
m_f	$\pm 20\ \mu$	0	$\pm 30\ \mu$	0	$\pm 30\ \mu$	0
m_ϕ	$\pm 0.49^c$	$\pm 0.45^c$	$\pm 1.60^c$	$\pm 1.58^c$	$\pm 1.12^c$	$\pm 1.12^c$
m_κ	$\pm 0.28^c$	$\pm 0.21^c$	$\pm 1.19^c$	$\pm 0.69^c$	$\pm 0.49^c$	$\pm 0.49^c$
m_ω	$\pm 0.71^c$	$\pm 0.70^c$	$\pm 2.36^c$	$\pm 2.35^c$	$\pm 1.07^c$	$\pm 1.07^c$
m_{b_x}	± 0.055 mm	± 0.013 mm	± 0.152 mm	± 0.044 mm	± 0.070 mm	± 0.026 mm
m_{b_y}	± 0.077 mm	± 0.053 mm	± 0.332 mm	± 0.175 mm	± 0.334 mm	± 0.075 mm

in which $f_R = f_L$, $m_{f_R} = m_{f_L}$, etc. analogously. These are to be combined with the weight coefficients as determined by Hallert, $Q_{\phi\phi} = h^2/b^2d^2$, to obtain accuracies of exterior orientation elements after orientation, for example:

$$m_\phi^2 = m_0^2(Q_{\phi\phi} + Q_{\phi\phi}') = m_0^2\left(\frac{h^2}{b^2d^2} + \frac{2\sin^4\phi}{f^2}\frac{m_{dx}'^2}{m_0^2} + \frac{2\sin^2\phi\cos^2\phi}{f^2}\frac{m_f^2}{m_0^2}\right)$$

Analogously one obtains:

$$m_\omega^2 = m_0^2(Q_{\omega\omega} + Q_{\omega\omega}') = m_0\left[\frac{3h^2}{4d^2}\cos^2\phi + \frac{2}{3b^2}\sin^2\phi + \frac{2\sin^4\phi}{f^2}\frac{m_{dy}'^2}{m_0^2}\right]$$

$$m_\kappa^2 = m_0^2(Q_{\kappa\kappa} + Q_{\kappa\kappa}') = m_0\left[\frac{3h^2}{4d^2}\sin^2\phi + \frac{2}{3b^2}\cos^2\phi + \frac{2\sin^2\phi\cos^2\phi}{f^2}\frac{m_{dx}'^2}{m_0^2}\right]$$

$$m_{b_x}^2 = m_0^2(Q_{b_xb_x} + Q_{b_xb_x}') = m_0\left[\frac{h^2}{2d^2} + \frac{\sin^2\phi}{f^2}(b^2\sin^2\phi + 2bh\sin\phi\cos\phi + 2h^2\cos^2\phi)\right.$$
$$\left.\frac{m_{dx}'^2}{m_0^2} + \frac{\cos^2\phi}{f^2}(b^2\sin^2\phi - 2bh\sin\phi\cos\phi + 2h^2\cos^2\phi)\frac{m_f^2}{m_0^2}\right]$$

$$m_{b_y}^2 = m_0^2(Q_{b_yb_y} + Q_{b_yb_y}') = m_0\left[\frac{3h^2}{4d^2} + \frac{h^2}{d^2} + \frac{2}{3} + \frac{2h^2\cos^2\phi}{f^2}\frac{m_{dy}'^2}{m_0^2}\right]$$

Table 2 shows a numerical example of the standard errors of the orientation elements for normal-angle convergent photography and wide-angle vertical photography under different assumptions of interior orientation accuracy; m_ϕ, m_κ, m_ω, m_{b_x} and m_{b_y} are computed:

(1) for the case of careful elimination of interior orientation errors (columns 1 and 2)

(2) for the case of ordinary plotting procedures, without exercising special care (columns 3 to 6).

The calculations are made for convergent photography (columns 1 to 4) as well as for vertical photography (columns 5 and 6). Note that m_f is larger for convergent photography as compared to vertical photography since two cameras and films are used.

Columns 1, 3 and 5 give orientation errors including error sources of interior orientation. These are not considered in columns 2, 4 and 6, as obtained directly by Hallert's coefficients. The differences between columns 1 and 2, 3 and 4, 5 and 6 represent the influence of interior orientation errors on relative orientation.

OFEK07686

The comparison shows that the influence of interior orientation errors on relative orientation remains negligible as long as these remain very small; if they reach amounts of 100μ, as it may happen easily if the diapositives are inserted into the plate carriers by eye, relative orientation is disturbed mainly in s, by, and bz.

It is therefore essential that:

(a) an accurate camera calibration of the convergent cameras will ensure the position of the principal-point and the camera constant to $\pm 20\mu$ or better.

(b) that the diapositives are carefully placed into the plate holders of a carefully calibrated stereoplotter by means of magnifying glasses.

(c) that a film shrinkage correction according to known and measured fiducial mark distances $f=(a'-a)f/a$ is made from photograph to photograph.

Interior Orientation Errors and Aerial Triangulation

Further considerations have to be made, if bridging procedures should be applied to convergent photography.

When using the zero base model to join forward and rear pictures of the camera couple of one exposure station, the orientation formulae become:

$$db z_L = \frac{h}{2d}(py_3 - py_4);$$

$$h d\phi_L + db x_1 = \frac{1}{3}(px_1 + px_3 + px_5);$$

$$db y_L = \frac{h^2}{2d^2}(py_3 - py_1 + py_5 - py_2) - py_1;$$

$$dx_L = (px_3 - px_4)\frac{\cos\phi}{2d} - (py_3 - py_1 + py_5 - py_2)\frac{h\sin\phi}{2d^2};$$

$$d\omega_L = -(px_3 - px_4)\frac{\sin\phi}{2d} - (py_3 - py_1 + py_5 - py_2)\frac{h\cos\phi}{2d^2};$$

by analogous compensation of x and y parallaxes at points 1 to 5 by interior orientation errors we obtain:

$$db z_L = h\cos^2\phi\left(\frac{df_L}{f_L} - \frac{df_R}{f_R}\right) - h\sin\phi\cos\phi\left(\frac{dx_L'}{f_L} - \frac{dx_R'}{f_R}\right);$$

$$h d\phi_L + db x_1 = h\cos^2\phi\left(\frac{dx_L'}{f_L} - \frac{dx_R'}{f_R}\right) + h\sin\phi\cos\phi\left(\frac{df_L}{f_L} - \frac{df_R}{f}\right);$$

$$db y_L = h\cos\phi\left(\frac{dy_R'}{f_R} - \frac{dy_L'}{f_L}\right)$$

$$dx_L = d\omega_L = 0$$

It becomes evident that:

(a) the theoretical condition to keep $by_L = by_R$, $bz_L = bz_R$, $bx = b_{\text{constant}}$ (Ackermann 1956) should not be maintained in the plotter; it is important to remove all x and y parallaxes in the zero base model, and this, on account of interior orientation errors, will only be possible by using by, bz, and bx, and therefore we obtain the general statement: $by_L \neq by_R$, $bz_L \neq bz_R$ and $bx \neq b_{\text{constant}}$.

(b) bx and ϕ cannot be determined independently. It is common procedure to set $bx = 0$ (or b_{constant}) and to move ϕ until the x parallaxes become zero. This

OFEK07687

is incorrect since δx should correspond to the unknown interior orientation error. It is therefore best to determine the angle of convergence γ out of the mean between several preliminary zero base model orientations and to keep the difference in ϕ consistently equal to γ for each zero base orientation while bridging. A systematic error in γ will not harm the triangulation result.

(c) such a procedure will also hold, if errors of synchronization between the two cameras are present. This would still tolerate errors of up to 1/75 sec. (allowing for up to 0.6°/sec. angular velocity of the aircraft), even though most convergent cameras are usually synchronized to $\pm 1/500$ sec.

Other Factors Limiting Photogrammetric Accuracy

Distortion and interior orientation are inseparable according to the physical definition of the optical-axis as the axis of symmetry. Distortion is usually determined from the point of symmetry. Using compensation devices in plotting, a maximum of $\pm 4\mu$ distortion difference is encountered (Hothmer, 1959, pp. 60–78).

Its components $\Delta x'$ and $\Delta y'$ in the photo act as Δx and Δy in the model:

$$\Delta x = \frac{h \cdot f}{(z \cos\phi - x' \sin\phi)^2} \Delta x'$$

$$\Delta y = \frac{hy \sin\phi}{(f \cos\phi - x' \sin\phi)^2} \Delta x' + \frac{h}{f \cos\phi - x \sin\phi} \Delta y'$$

It can be shown that a radial symmetrical distortion has a more favorable effect on convergent than on vertical models. Since in convergent photography two different cameras with different actual distortions may be in use, this may not hold true, however, and it is of prime importance to use cameras in the convergent mount that have been particularly matched or calibrated together.

Film errors occur during the exposure and during the processing and aging of the exposed material. It is well established that these constitute a major part of photogrammetric inaccuracies. In this respect, a smaller angular field is more favorable than a wider one, for two reasons:

First, the incoming rays at the time of the exposure will hit the emulsion (which is only flat to $\pm 20\mu$) at a lesser angle ($\Delta r = r \tan\omega$); second, film errors will be enlarged more for wider angle photography because of the bigger magnification ratio. It is immaterial whether vertical or convergent photography of the same angular field is used.

It should be noted, however, that the film of both cameras should be treated similarly in processing; otherwise different regular film shrinkage, particularly different amounts of differential film shrinkage may result.

Instrument errors. Plotter accuracy is usually independent on whether vertical or tilted photographs are used. It is of importance however that the plotter is carefully calibrated for its interior orientation when convergent photography should be plotted.

The theoretical increased accuracy of convergent photography over vertical photos will be reached only when the above precautions are taken.

Practical Point and Bridging Accuracy

That this is possible was shown by a practical test:

Simultaneously exposed normal angle convergent photography taken by a Zeiss 2× RMK 21/18 was compared with wide-angle RMK 15/23 photography assuring the same conditions for weather, illumination, film treatment and control. In one strip of 15 and in two strips of 11 models each 278 control points were available. The photography was taken from 15,000 feet. The evalnation was done on a C8.

OFEK07688

TABLE 3

ACCURACY COMPARISON FOR SIMULTANEOUS VERTICAL AND CONVERGENT PHOTOGRAPHY

	m_P	m_H
Wide-Angle Vertical: points:	$\pm 0.15^{o}/_{oo}h$	$\pm 0.12^{o}/_{oo}h$
Normal-Angle Convergent: points:	$\pm 0.10^{o}/_{oo}h$	$\pm 0.08^{o}/_{oo}h$
Wide-Angle Vertical: bridging 15 models	$\pm 0.26^{o}/_{oo}h$	$\pm 0.36^{o}/_{oo}h$
Normal-Angle Convergent: bridging 15 models	$\pm 0.24^{o}/_{oo}h$	$\pm 0.27^{o}/_{oo}h$

The results of the test are considerably better than any values published thus far. (See Table 3).

Bridging, which can be considered as one of the most crucial tests for the accuracy of a photogrammetric procedure, was successful both in horizontal and vertical accuracy with higher accuracy than corresponding vertical procedures with up to 15 models.

Figure 3 demonstrates position and vertical accuracy versus the number of models bridged for normal-angle convergent and wide-angle vertical photographs.

This indicates a more unfavorable error propagation for convergent photography, which can be found by theory as well, by adding the influences of model orientation with those of the zero base model:

$$\Delta\phi_n = \sum_1^n d\phi_{\text{Model}} + \sum_0^n d\phi_{\text{Zero Base Model}}$$

$$\Delta\xi_n = \left(\cos\phi + \frac{k}{b}\sin\phi\right)\left(\sum_1^n d\omega_{\text{Model}} + \sum_0^n d\omega_{\text{Zero Base Model}}\right)$$

$$\Delta\xi_n = \left(\cos\phi - \frac{k}{b}\sin\phi\right)\left(\sum_1^n dx_{\text{Model}} + \sum_0^n dx_{\text{Zero Base Model}}\right)$$

$$\Delta bz_n = -b\sum_1^n (n-i)d\phi_{\text{Model}} - b\sum_1^n (n-i)d\phi_{\text{Zero Base Model}} + \sum_1^n dbz_{\text{Model}} + \sum_0^n dbz_{\text{Zero Base Model}}$$

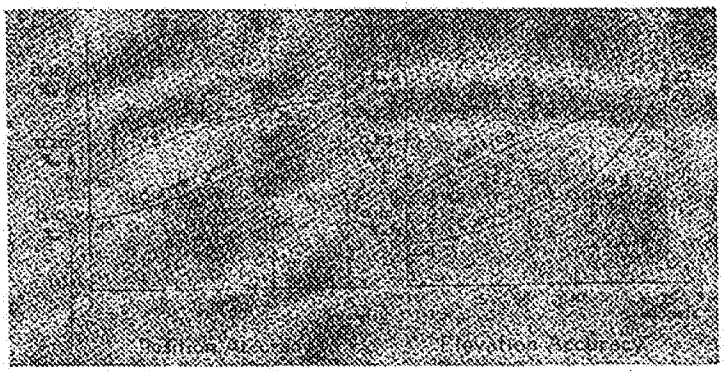


FIG. 3. Bridging Accuracy as a Function of the Number of Models for Vertical and Convergent Photographs.

OFEK07689

$$\Delta by_n = \left(\cos\phi - \frac{h}{b}\sin\phi\right)\cdot b\cdot\left[\sum_1^n (n-i)\,dx_{\text{Model}} + \sum_1^n (n-i)\,dx_{\text{Zero Base Model}}\right]$$

$$\Delta X_n = -\frac{b}{h}\left\{\sum_1^n (n-i+1)\,dz_{i-1} + h\sum_1^n\left[\frac{h}{b}\left(1+\frac{b^2}{h^2}\right)+\frac{b}{h}(n-1)\right]\cdot(d\phi_{\text{Model}} + d\phi_{i-1\text{Zero Base Model}}) + \sum_1^n (2n-2i+1)(dbz_{\text{Model}} + dbz_{i-1\text{Zero Base Model}})\right\}$$

For 15 models the ratios between the error propagations of normal-angle convergent and wide-angle vertical photography become

for the azimuth 6.1,
for warping of the strip in flight direction 3.6,
for warping of the strip perpendicular to flight direction 0.6,
for scale transfer 5.8.

Considering that for a certain strip length the number of models varies with the type of camera used, this is not nearly as bad as it looks. (See Table 4.)

The test proves that aerial triangulation with convergent photography is feasible, if properly executed.

Analytical Procedures

Analytical procedures are favorable to the use of convergent photography, particularly if signalized points are used and the photo co-ordinates are measured on a mono-comparator.

It is then that difficulties due to centering of the diapositives, a main source of interior orientation errors, can be avoided, since the fiducial marks can be more precisely read then centered.

Even if calibration errors of a certain camera assembly should be present, these together with eventual distortions can be eliminated analytically, after the camera has taken photographs over a known test area.

An investigation of this nature is currently being made and the results are expected as soon as our new computer installation at U.N.B. becomes operational.

It is up to the individual to decide whether it is of advantage or not to employ the more awkward convergent photography. It is more important, however, to show that theory and practice can be brought into accordance when using proper procedures.

Table 4

Number of Models for Equal Strip Length and Different Types of Photography

	Based on Equal Photo Scale	*Based on Equal Flying Height*
Vertical Normal Angle	100	100
Vertical Wide Angle	68	50
Vertical Super Wide Angle	70	29
Convergent Normal Angle	54	52
Convergent Wide Angle	41	24

References

Ackermann, F. "Zur rechnerischen Orientierung von Konvergentaufnahmen." Bildmessung und Luftbildwesen, 1956, pp. 24–37.

Finsterwalder, S. "Eine Grundaufgabe der Photogrammetrie und ihre Anwendung auf Ballonaufnahmen." Abh. d.K. Bayer. Akad. d. Wissensch, II. Kl., XXII Bd., II. Abt., pp. 225–260, 1903.

Griffin, E. P. "Vertical Photography for Aerotriangulation." Photogrammetric Engineering, 1959.

Hallert, B. "Some remarks concerning the theory of errors for convergent aerial pictures in comparison with near vertical pictures. Photogrammetric Engineering 1954, pp. 749–757.

Hothmer, J. "Possibilities and Limitations For Elimination of Distortion in Aerial Photographs." *Photogr. Record*, 1958. II/12 pp. 426–445, 1959. III/13, pp. 60–78.

Konecny, G. "Aerotriangulation mit Konvergentaufnahmen." Veröff d. Deutschen Geodatischen Kommission, Reihe C., Heft 47, Munich 1962.

Pennington, J. T. "Aerotriangulation With Convergent Photography." Photogrammetric Engineering, 1954, pp. 76–81.

Theis, J. B., "Increased Base: Height Ratio." Photogrammetric Engineering, 1958, pp. 127–132.

OFEK07690

Manual
of
Photogrammetry

Third Edition

In two volumes
VOLUME II



Editor-in-Chief
Morris M. Thompson

Associate Editors
Robert C. Eller
William A. Radlinski
Julius L. Speert

American Society of Photogrammetry

OFEK07691

© *Copyright 1944, 1952, 1966 by the* AMERICAN SOCIETY OF PHOTOGRAMMETRY

All rights reserved. This book, or parts thereof, may not be reproduced in any form without permission of the publishers.

Third Edition

Library of Congress Catalog Card Number: 65-20813

PUBLISHED BY

AMERICAN SOCIETY OF PHOTOGRAMMETRY

6269 Leesburg Pike

Falls Church, Va. 22044

Printed in the United States of America by

GEORGE BANTA CO., MENASHA, WISCONSIN

OFEK07692

Airborne Video for Near-Real-Time Vegetation Mapping

By Lee A. Graham

One difficult, time-consuming, and expensive component of many natural resource mapping and monitoring efforts is obtaining adequate, geographically unbiased, verification and validation data. During the last one and a half years, Arizona GAP Analysis project investigators have been generating a statewide vegetation map from Landsat Thematic Mapper (TM) imagery. The GAP project is part of a nationwide effort to obtain biodiversity knowledge for all areas managed exclusively or primarily for long-term maintenance of native species populations and natural ecosystems (Scott et al. 1993).

The GAP process relies on a map of actual vegetation, supplemented by vertebrate species locations, to predict individual species distributions and ultimately species richness; thus the vegetation map must be recent and of good quality. This is a complicated process in Arizona, since the vegetation on most of the tribal lands has not been mapped and because many areas cannot easily be accessed from the ground. Also, many vegetative communities contain sparsely distributed and/or small shrub and scrub species growing on widely varied soils that cannot be readily differentiated with TM-band spectral properties. Consequently, we needed to find a way to verify the information for this mapping effort within a very short timeframe. The airborne video acquisition, interpretation, and ground-truthing procedures described here were developed to meet this need. A central component of the GAP project is a video data acquisition system that uniquely tags every video frame with global positioning system (GPS) data instead of SMPTE (Society for Motion Picture and Television Engineers) longitudinal time code.

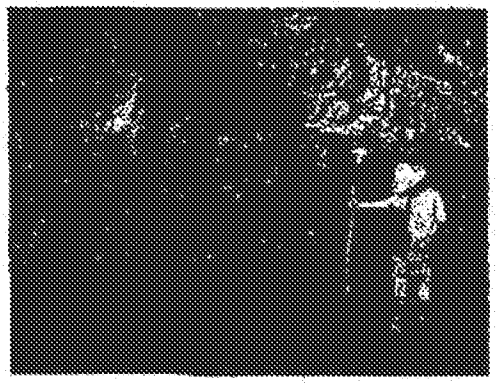
Background

Airborne videography is an inexpensive and rapid method to acquire images for resource evaluation and mapping. Natural-color and multispectral airborne video imagery can assess and monitor grassland and rangeland vegetation, crop conditions, forest pest outbreaks, soils, water quality, and hazardous waste sites (Nixon et al. 1985, Everitt et al. 1986, Everitt et al. 1989, Lee 1990, Stutte and Stutte 1990, Marsh et al. 1991, Myhre et al. 1991). Video data have also characterized reflectance patterns for some tree species (Everitt and Nixon 1985, Vlcek and King 1984) and physiological activity and biomass productivity in mostly agricultural settings (Everitt et al. 1986, Marsh et al. 1990, Stutte and Stutte 1990).

Video images are generally integrated with geographic position information by using caption generators, which encode Loran-C or GPS positions as text directly on the video frame (Myhre et al. 1991, Bobbe 1992, Evans 1992). However, these data cannot be retrieved automati-

Mapping Plant Species

The Portland USDA Forest Service Engineering Department has been working on an ambitious project to bring together a variety of forest maps and data into a master composite map covering a 24-million-acre tract. Presenting the data in a series of layers allows resource managers to see the relation among the biological, physical, and human systems that affect the life of the forest.



John Good, Trimble Navigation

To ensure that each map layer registers accurately with the others, GPS survey receivers establish reference points in the field. Then, GPS in conjunction with aerial photography creates a digital elevation model of the site. Soil types, vegetation species, wildlife habitats, roads, and structures are all plotted on this base map. GPS mapping receivers are used to locate items that do not show up on the photographs, such as archaeological sites and wilderness boundaries.

Farther south, in the Chiapas region of southern Mexico, researchers have used GPS to map the distribution of plants. More than 1,500 species are purported to have medicinal value to native inhabitants. This study will determine which plants actually have medicinal value and map their locations.

—*Paul Roberts, Trimble Navigation*

OFEK07768

cally by a computer. GPS caption data must be manually matched to the digital GPS data file logged by the GPS recording device (generally a polycorder or computer running software like GeoLink). GPS typically updates on a 1-second interval, so the same caption is displayed for each of the 30 frames in one second of video, complicating the matching process.

SMPTE time code is a standard audible timing signal that identifies each frame of a television picture (Horita, Inc. 1991). Rather than being recorded as part of the picture, it is recorded on an audio or special time code channel of the VCR. A time code generator can produce a time code while the picture is being recorded. A time code reader converts the audio signal back to digital information, which can be displayed by a VCR time code counter or a computer running a program that displays time code. A time code window-inserter then can function as a caption generator, dubbing time code values on the picture during playback. Horita, Inc., has engineered products that let GPS receivers substitute highly accurate GPS time in place of SMPTE signals. As a result, each frame is uniquely labeled with computer-retrievable GPS time, providing a way to precisely and automatically match video-recorded GPS time with the position information in the GPS data file.

Integrated GPS/Video System

Figure 1 shows how a single camera was integrated with an airborne GPS/video system for the Arizona GAP project. Although Trimble Pathfinder GPS receivers were used, the Horita unit and GeoLink software can work with other receivers.

The GAP application used a single natural-color camera; however, a second natural-color camera, as well as thermal or infrared cameras, can enhance the system's usefulness for riparian and wetland mapping and for monitoring fire lines. A simple Y-cable allows simultaneous recording of GPS-derived time codes to correlate frames from multiple tapes.

Interpretation Station

An interpretation station *(fig. 2)* permits simultaneous viewing of video imagery and retrieval of the corresponding GPS time data by a microcomputer. GPS times can be "logged" using computer programs provided with the time code units. PC-LOG can record GPS time data for frames of interest as well as separate comment fields in a simple text format. TC-CALC allows the user to define a keyboard sequence ("hot key") that inserts GPS time into any application—for example, a field in a multiple-attribute dBase application. A complete log file is then matched to the GPS (or Apollo Loran-C) data file with an in-house program (the GAP project used GMATCH) that computes the best geographic position for each logged frame—after screening for position data reliability. GMATCH can produce two basic outputs: a text file that reports GPS time, estimated position, aircraft elevation, and comments or database information; and geographic information system (GIS) point data files and multiple-attribute tables that record the above information with flight and mission data (useful for project management). Both GMATCH products can be generated, printed, and input to a GIS within 5 to 10 minutes of interpreting a tape.

A user may want to determine the geographic position of a frame along a transect without actually logging a tape. Horita accessories can insert and display information on latitude, longitude, and

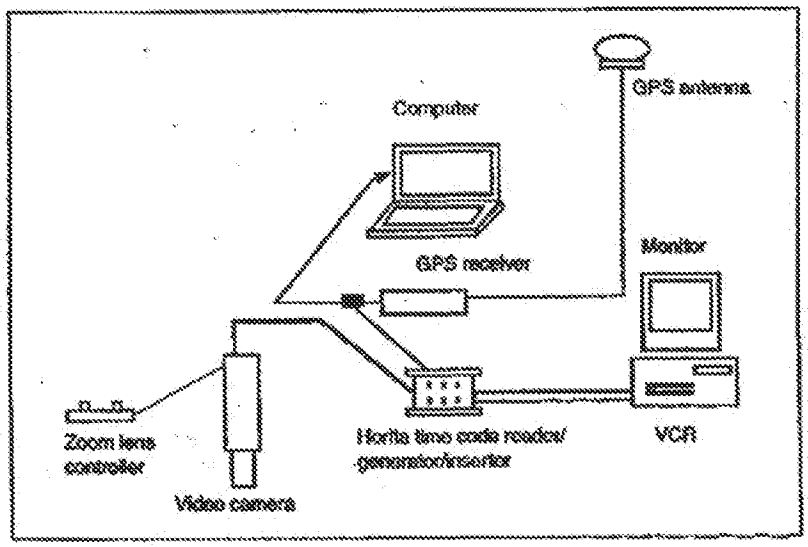


Figure 1. Single-camera integrated airborne GPS/video data acquisition system.

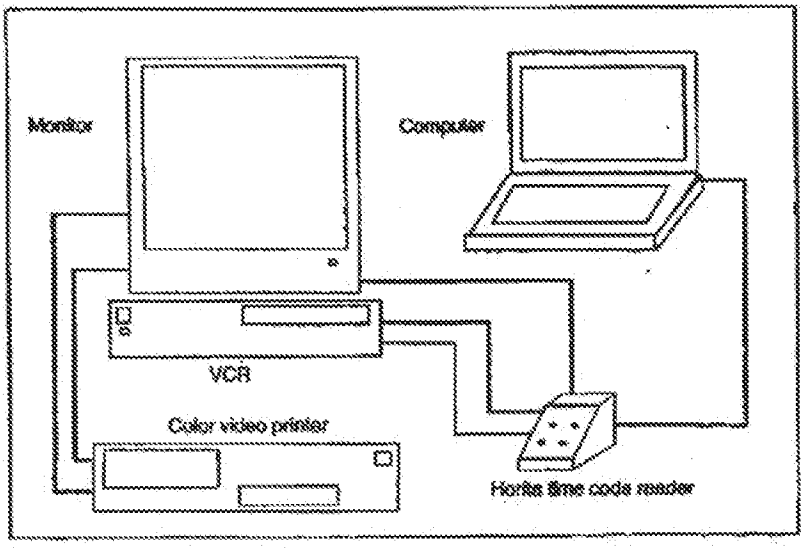


Figure 2. Single-VCR integrated GPS/video data interpretation station.

OFEK07769

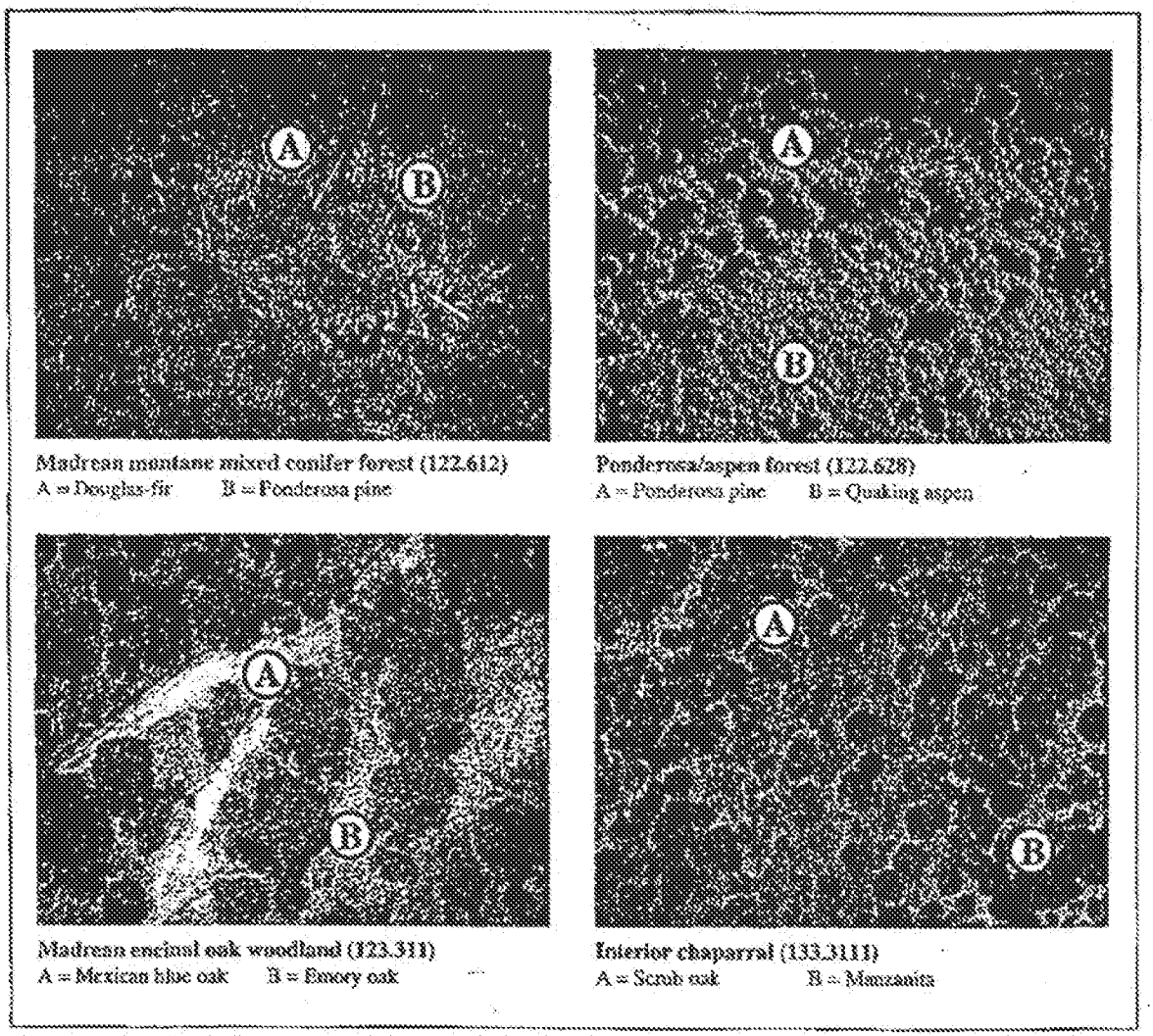


Figure 3. Species differentiation using natural-color airborne video data of Arizona forest, woodland, and chaparral communities (numbers correspond to Brown, Lowe, and Pase vegetation association codes).

elevation data at 1-second intervals. However, for the GAP project a program was developed to convert GPS data files directly to GIS files and multiple-attribute tables. Once imported to GIS format, these files can highlight locations along a video flight path, with the GPS-generated time code dubbed on every frame. ArcView, a desktop mapping system, proved to be a powerful tool for viewing and querying GIS files produced from both GPS and video interpretations. It also provided an easy way to integrate data for visual and analytical inspection with other GIS and image data.

A color video printer can produce prints or transparencies of video frames directly from the tape. While video prints vary greatly in print quality and image size, the GAP project model could produce 3-inch by 4-inch color prints in approximately 1 minute at a cost of about 45 cents per print.

Study Methods

Our goal was to obtain a 3 percent sampling density of the state. Some areas were not sampled due to flight restrictions (Grand Canyon, military areas) or lack of relevance to GAP goals (metropolitan Phoenix). The project effectively broke out into southern and northern Arizona phases, which offered different ways to achieve a representative vegetation point sample. The southern phase, acquired between November 1991 and May 1992 and covering about one-third of the state, used a 1970s set of vegetation maps developed by the Arizona Game and Fish Department to locate transects so that the widest array of vegetation communities could be sampled. However, these maps were found to be unreliable for desert, grassland, and woodland areas. Thus the remaining two-thirds of the state used a 30-kilometer grid of transects. The southern phase was flown at 3,000 feet and took 12 days; the northern phase only took 10 days and was flown at 2,000 feet.

These flight elevations resulted in swath widths of approximately 1/3 to 1/2 mile. At approximately 9-second intervals, the camera lens was zoomed from

OFEK07770

9.5 mm (wide angle) to 143 mm (zoom), left down for approximately one to three seconds, then returned to the wide angle. This procedure provided almost continuous wide-angle coverage with punctuated zoomed subsamples. The ground distance between adjacent zoomed samples ranged from 3,980 to 4,980 feet. Wide-angle views provided information on landscape continuity and configuration, while mid- and full-zoom images were best to identify individual species and relative ground cover.

A straightforward and inexpensive technique was developed to recognize species and community types from airborne video imagery. During visits to all representative plant communities, dominant and codominant species were identified and quantified according to prominence and relative ground cover. Common associate species were also noted. Series-level names under the Brown, Lowe, and Pase vegetation classification system (Brown 1982) were determined by the dominant and codominant species, while association names (the next more detailed level in the hierarchy) were also based on nondominant but common species.

GPS data files for each mission were converted to GIS data files so that position data could immediately be accessed and queried using MIPS (Map and Image Processing System by MicroImages, Inc.) and later Arcview. Tapes were then viewed, and GPS times of representative and accessible vegetation communities were logged along with a short vegetation description. Wide-angle and zoomed video prints were made for logged points. The GMATCH program computed geographic coordinates for each set of prints. GIS data files were also produced so that printed hard copies could show locations of logged points overlaid on a scanned and georeferenced road map. A large map was also pieced together from the separate map prints to provide an overview of logged sample points for the entire northern Arizona study area. Altogether, 52 hours of video from northern Arizona were interpreted within one month of data acquisition.

The overview map was used to plan a route to visit the greatest number of field points within seven days. Geographic coordinates for each day were entered into a Trimble Basic GPS unit as waypoints, and the GPS unit and remote antenna were installed in the vehicle. Once in range of the site, the GPS unit was dismounted and used to navigate the crew to the field site. The GPS information and color video prints permitted precise location of the areas. Field data sheets logged plant species, prominence, and relative cover for each frame, along with observations on land form, soil, elevation, slope, and aspect. Video prints were annotated to illustrate species identifications, along with camera photo-points and angles from ground photographs. Samples of unknown species were collected and additional GPS-referenced vegetation observations recorded as the vehicle traveled between sites, providing yet another source of verification and validation data.

—All the field information was later organized by series and used as photointerpretation keys for assigning Brown, Lowe, and Pase series codes to video point samples. To assign vegetation codes, the interpreter had to be able to recognize and assess relative ground cover and/or prominence of species within a zoom or set of zooms.

Figure 3 shows zoomed images from selected communities in Arizona and identifies dominant and codominant species for each community. Video imagery proved suitable for identifying these species, along with many associate species within a frame. It was therefore possible to identify most airborne video points to an association or even subassociation level of the Brown, Lowe, and Pase system. Texture, shape, land form, and cultural features often proved to be better guides for identification than color, which varied according to sun angle, topography, and atmospheric conditions.

The video point-sampling technique was particularly effective for forest, woodland, shrub, and desert scrub communities. It was less effective for differentiating grassland types. Semidesert tobosa grass, sacaton, and bunchgrass and grass-scrub series can be easily identified. However, in other areas it was impossible to assign a plains grassland area to the grama or blue-stem/buffalo grass series. In these instances, information on soil type, land form, and topography helped the decision process.

Results

A total of 11,223 point samples were generated for the entire project—an average of 1,006 (standard deviation = 283.5)

Fertilizer Application

Weyerhaeuser Timberlands is presently using GPS for a variety of purposes: acreage determination, inventory, stream assessment and monitoring of fish habitat, and mapping new roads. Currently, the aviation department is testing "agricultural" systems that will track and record fertilizer applications.

The system is fairly simple. A receiver in the aircraft uses satellite signals as well as the differentially corrected signals broadcast (via radio) from the ground base station. The real-time coordinates are then displayed on a screen mounted in the aircraft. A lightbar connected to the system guides the pilot in keeping the craft in a straight line to prevent possible overlaps and gaps in the flight pattern. It will also log the exact coordinates when the tank runs out of fertilizer so that the pilot will know exactly where to begin applications again on that line. Although steep terrain may cause difficulties in satellite reception, such problems can be reduced or avoided by proper planning—such as determining the best windows of satellite availability and geometry.

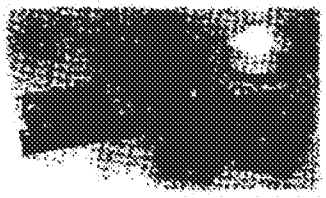
Using GPS to control fertilizer application is still in the testing stage, and Weyerhaeuser is in the process of evaluating the results. However, with a potential accuracy of greater than one meter, the system promises much improved cost effectiveness, along with the traditional GPS advantages of faster data acquisition, very accurate position information, compatibility with most geographic information systems, and the ability to tie locations to known coordinates.

—Diana Broome, GPS coordinator, Weyerhaeuser Timberlands, Federal Way, Washington.

OPEK07771

FOUR WAYS
TO LINK VIDEO IMAGES
TO GPS DATA

GPS51 - generates industry standard IRIG/SMPTE time code synchronized to GPS time-data from your receiver. SMPTE time code recorded on an audio channel or time code track provides a permanent and direct time-index link between recorded video images and the GPS location stored in your GPS data logger. $625



GPS52 - adds continuous multiplexing of GPS position data into the user-bits of the SMPTE time code to log position and velocity data directly onto the video tape. $1,397

GPS53 - includes the functions of GPS51 and GPS52 and provides for multiplexing of additional analog and serial data inputs, such as temperature, heading, etc., into the SMPTE time code. $1,850

GPS-OSD - GPS video title adds real-time GPS position information to your video recording. Interfaces directly with your GPS receiver and video recorder to add onscreen titling as well as GPS data. $995

Horita
P.O. Box 3900, Mission Viejo, CA 92690
(714) 488-0240

GIS Applications In Natural Resources

Selected Papers from the World's Foremost GIS/Natural Resources Conference Series. Edited by Michael Heit & Art Shortreid. Foreword by Charles E. Hamphreys

This 400-page, hard-cover book is a compilation of the best papers given at the GIS'87 - GIS'91 conferences sponsored by Forestry Canada. A comprehensive overview of GIS applications in natural resources in the United States, Canada and Europe, the book includes articles on forestry, wildlife, timber inventory, harvest scheduling, environmental assessment, data sharing, hardware, training, data conversion, remote sensing and more. Through the cooperative efforts of the editors and GIS World, Inc., we are pleased to be able to offer GIS Applications in Natural Resources at far less than its market price:

$29.95 (U.S.) and $35.95 (Canada). MasterCard, VISA, purchase orders or checks accepted. Add $3.00 for overseas air mail. Order toll free from the U.S. or Canada

1-800-GIS-WRLD

GIS World, Inc.

155 E. Boardwalk Drive, Suite 350, Fort Collins, CO 80525, USA
(303) 223-4848 • FAX (303) 224-0786
anywhere

per scene—and 142 vegetation series, associations, and/or subassociations were identified during this process.

It was not practical to differentially correct GPS data for this scale and time period. The project was begun before many base stations were installed throughout the state, and those installed were not always in operation. An informal evaluation of the accuracy of GPS locations was determined for 19 frames containing easily identified cultural features. A comparison of GPS positions against coordinates indicated that they fell within 0 to 484 feet of the locations provided from the maps. This estimate appeared reasonable, considering selective availability and aircraft shifts due to turbulence. Since the minimum mapping unit mandated for GAP was 247 acres, this level of accuracy was deemed acceptable.

The video point-sample database provided verification and validation information for the Arizona GAP vegetation map. Two-thirds of the video point samples were used to merge and label classes that result from an unsupervised classification of band ratios based on Landsat bands 3, 4, and 5. The remaining points were coordinated with GPS-referenced road-log observations to validate the final map.

Summary

The Arizona GAP project used data acquisition, interpretation, and ground-truthing procedures that relied on integrated airborne GPS/video systems. These techniques allowed the collection of thousands of vegetation point samples throughout the state. The samples were less geographically biased than those from traditional road-log verification and validation ventures. The point-sample database was used with Landsat TM imagery to build a series-level map of actual vegetation. Extensive documentation of plant species and communities at multiple scales as they appear on natural-color video imagery were organized into photo-interpretation keys, which will prove useful for subsequent mapping and monitoring efforts.

Literature Cited

Bossler, T. 1992. Real-time differential GPS for aerial surveying and remote sensing. GPS World 3:18–22.

Brown, D. 1982. Biotic communities of the American Southwest—United States and Mexico. Desert Plants 4:1–342.

Evans, D.L. 1992. Using GPS to evaluate aerial video missions. GPS World 3:24–29.

Everitt, J.H., M.A. Alaniz, and M.R. Davis. 1989. Using multispectral video imagery for detecting soil surface conditions. Photogramm. Eng. & Remote Sens. 55:667–71.

Everitt, J.H., M.A. Hunter, D.E. Escobar, R.R. Nixon, and B. Pinkerton. 1986. Assessment of grassland phytomass with airborne video imagery. Remote Sens. of the Environ. 20:299–306.

Everitt, J.H., and R.R. Nixon. 1985. False color video imagery: a potential remote sensing tool for range management. Photogramm. Eng. & Remote Sens. 51:875–79.

Horita, Inc. 1991. Horita TBG-50 SMPTE time-code reader, generator, window inserter—user manual. Horita, Inc., Mission Viejo, CA.

Lee, C.T. 1990. The use of multispectral video remote sensing to monitor suspended sediment concentrations. PhD thesis, Univ. Ariz., Tucson.

Maxell, S.E., J.L. Walsh, C.T. Lee, and L.A. Graham. 1991. Multitemporal analysis of hazardous waste sites through the use of a new bi-spectral video remote sensing system and standard CIR photography. Photogramm. Eng. & Remote Sens. 57:1221–26.

Maxell, S., J.L. Walsh, and C.F. Hutchinson. 1990. Development of an agricultural land-use GIS for Senegal derived from multispectral video and photographic data. Photogramm. Eng. & Remote Sens. 56:351–57.

Meisner, R.J., L.A. Graham, and C.W. Shoffner. 1991. Airborne video technology: protecting natural resources with remote sensing. *In* Proceedings, third USDA Forest Service remote sensing applications conference, p. 293–98. Am. Soc. of Photogramm. & Remote Sens., Bethesda, MD.

Nixon, P.R., D.E. Escobar, and R.M. Menges. 1985. Use of multiband video systems for quick assessment of vegetal condition and discrimination of plant species. Remote Sens. of the Environ. 17:203–8.

Scott, J.M., F. Davis, B. Csuti, R. Noss, B. Butterfield, C. Groves, H. Anderson, S. Caicco, F. D'Erchia, T. Edwards, J. Ulliman, and R. Wright. 1993. GAP analysis: a geographic approach to protection of biological diversity. Wildl. Monogr. 123:1–41.

Stuth, C.A., and G.W. Stuth. 1990. Qualification of field crop stress areas using videography. *In* Proceedings, 12th biennial workshop on color aerial photography and videography, p. 88–93. Am. Soc. for Photogramm. & Remote Sens., Bethesda, MD.

Vicen, J., and D. King. 1986. Digital analysis of multispectral video imagery. *In* Technical papers, 50th annual meeting of the American Society of Photogrammetry, p. 628–32. Am. Soc. for Photogramm., Washington, DC.

ABOUT THE AUTHOR

Lee A. Graham is assistant professor, School of Renewable Natural Resources, University of Arizona, Tucson.

OFEK07772

INT. J. REMOTE SENSING, 2002, VOL. 23, NO. 7, 1527–1556

Taylor & Francis
Taylor & Francis Group

Geocoded digital videography for validation of land cover mapping in the Amazon basin

L. L. HESS[1]*, E. M. L. M. NOVO[2], D. M. SLAYMAKER[3], J. HOLT[4]†, C. STEFFEN[2], D. M. VALERIANO[2], L. A. K. MERTES[5], T. KRUG[2], J. M. MELACK[1], M. GASTIL[1], C. HOLMES[3] and C. HAYWARD[3]

[1]Institute for Computational Earth System Science, University of California, Santa Barbara, California 93106, USA
[2]National Institute for Space Research (INPE), P.O. Box 515, São José dos Campos, SP, 12201-970, Brazil
[3]Department of Natural Resource Conservation, University of Massachusetts, Amherst, Massachusetts 01003, USA
[4]Jet Propulsion Laboratory, MS 300–243, 4800 Oak Grove Drive, Pasadena, California 91109-8099, USA
[5]Department of Geography, University of California, Santa Barbara, California 93106, USA

Abstract. Four Validation Overflights for Amazon Mosaics (VOAM) aerial video surveys have been carried out in the Brazilian Amazon to provide ground verification for mapping of wetland cover with the Global Rain Forest Mapping (GRFM) Project JERS-1 (Japanese Earth Remote Sensing Satellite) mosaics of the Amazon basin. Surveys in 1995 and 1996, acquired with handheld analog camcorders from small aircraft, were timed to imaging of the GRFM low- and high-water mosaics, and limited to within 600 km of Manaus. For the 1997 and 1999 flights, digital camcorder systems were installed in the Bandeirante survey plane operated by Brazil's National Institute for Space Research. The VOAM97 and VOAM99 surveys circumscribed the Brazilian Amazon, documenting ground conditions at resolutions on the order of 1 m (wide-angle format) and 10 cm (zoom format) for wetlands, forests, savannas, and human-impacted areas. Global Positioning System (GPS) information encoded on the video audio track was extracted by mosaicking software that automatically generates geocoded digital mosaics from video clips. On the 1999 survey, a laser altimeter recorded profiles of terrain and vegetation canopy heights. A validation dataset was compiled from the videography for a portion of the GRFM mosaics extending 6° by 4° in longitude and latitude, using randomly selected points along flight lines. Other applications of the VOAM videography include acquisition of ground control points for image geolocation, creation of a high-resolution geocoded mosaic of a forest study area, forest biomass estimation, and rapid assessment of fire damage. Geocoded digital videography provides a cost-effective means of compiling high-resolution validation datasets for land cover mapping in remote, cloud-covered regions.

*e-mail: lola@icess.ucsb.edu
†Present address: Institute for Geophysics, Campus Mail Code R2200, University of Texas, Austin, Texas 78712, USA

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2002 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160110092687

1. Introduction

The capacity to generate satellite-based land cover maps at global or continental scales has evolved rapidly over the past two decades, in response to advances in the ability to receive, process, calibrate, classify, store, and disseminate large volumes of image data (Cihlar 2000, Defries and Belward 2000). As the geographic extent of land cover datasets has expanded, the task of acquiring adequate ground verification for accuracy assessment has become increasingly difficult. Validation strategies that work well for local studies, such as interpretation of aerial photographs or use of existing high-resolution maps, become impractical or impossible at subcontinental or global scales. Instead, a statistically valid verification can be based on high-resolution satellite imagery of a set of randomly selected sample pixels, as in the use of Landsat Thematic Mapper (TM) and SPOT imagery for validation of the 1 km DISCover land cover dataset (Muchoney *et al.* 1999, Scepan 1999). Validation sets can also be developed using intensively studied sites representative of biome types, building libraries of multi-sensor, multi-scale data; this is the strategy behind the development of the EOS Land Validation Core Sites (Justice *et al.* 2000). Rather than being randomly selected, such sites are located to take advantage of long-term ground and aircraft-based measurements.

This paper describes a solution to a subcontinental-scale validation problem that could not be addressed using satellite imagery or intensive study sites: mapping Amazonian wetlands at high-and low-water stages using the Global Rain Forest Mapping (GRFM) Project JERS (Japanese Earth Remote Sensing Satellite) mosaics of the Amazon basin (Rosenqvist *et al.* 2000). Aerial videography provided a means of sampling the study area of more than five million km^2 at reasonable cost. The JERS-1 Amazon Multi-season Mapping Study (JAMMS) wetlands mapping, a component of the NASDA (National Space Development Agency of Japan) GRFM project and of NASA's Large-Scale Biosphere–Atmosphere Experiment in Amazônia (LBA) project, will provide the first accurate estimates of the seasonal extent of inundation and wetland vegetation in the Amazon basin.

Synthetic Aperture Radar (SAR) sensors are well suited for remote sensing of tropical wetlands, since they penetrate cloud cover and can detect flooding beneath vegetation canopies; the L-band frequency and HH (horizontal send, horizontal receive) polarization of the JERS-1 satellite are optimal for mapping of inundation in flooded forests, which are extensive in the Amazon basin (Hess *et al.* 1995). JERS-1 image acquisitions for the two GRFM Amazon mosaics (Chapman *et al.*, 2002, this issue) were timed to correspond to high-and low-water conditions for the Amazon River. The classification system used for the wetlands mapping (Hess 1999) is based on vegetation structure and inundation state, important functional parameters required for modelling methane generation and other biogeochemical processes in wetland ecosystems (Sahagian and Melack 1998).

The typical seasonal variation in river stage along the central Amazon floodplain near Manaus is about 10 m, with daily changes in the range of 10 cm during periods of peak change (Irion *et al.* 1997). Fluctuations in floodplain water levels can cause large changes in radar backscattering properties of woody vegetation, since varying amounts of trunks and canopies are submerged. Properties of aquatic macrophyte stands are also dynamic, as floating grasses grow and senesce rapidly (Junk and Piedade 1997). The key requirements for an Amazon wetlands validation dataset are therefore (1) ability to detect both vegetation structure and inundation, even beneath plant canopies, and (2) simultaneity with, or at least seasonal proximity to, the JERS-1 imaging dates.

Use of optical satellite imagery was not an option for the JAMMS wetlands validation because one of the important cover classes, flooded closed-canopy forest, cannot be distinguished from non-flooded forest with optical satellite sensors (Melack *et al.* 1994). For cover types that are detectable with optical data, such as open water or aquatic macrophytes, insufficient cloud-free scenes were available from sensors such as TM during the GRFM Amazon mosaic acquisition periods of September–November 1995 and May–June 1996, or during those months in other years. Although much of the Amazon basin has been mapped at a scale of 1:50 000 or 1:100 000, the thematic information on these maps is not adequate for verifying classifications of inundated area at high- and low-water stages. Nor was it possible to extrapolate ground information from intensive study sites, since the few existing sites cover a very small area and represent a narrow range of wetland types. Aerial photography was not practical because of the cost of film processing, and the lack of efficient means of geocoding and analysing photographic data without using an expensive photogrammetric camera system.

Beginning in the 1980s, aerial videography has been increasingly used for applications where the advantages of videography relative to photography, such as lower cost and immediate availability of data, outweigh the disadvantages, such as poorer spatial resolution and difficulty of analysis (Meisner 1986, Mausel *et al.* 1992). King (1995) provides a comprehensive review of the evolution of video sensors and applications through the early 1990s. Many aerial videography applications focus on measurement of transient phenomena such as wildlife populations (Sidle and Ziewitz 1990, Strong and Cowardin 1995) and pest infestations (Everitt *et al.* 1994); on mapping dynamic land features such as wetland plant communities (Jennings *et al.* 1992) and coastal landforms (Eleveld *et al.* 2000); or on land cover mapping in remote areas with limited existing aerial photography and poor infrastructure (Marsh *et al.* 1994). The US National Biological Service's Gap Analysis Program, which required low-cost, high-resolution, regional sampling of complex plant communities, was an impetus to the improvement of video survey methods in the early 1990s. An important advance was automatic labelling of video frames with Global Positioning System (GPS) coordinates, using a time-code generator (Graham 1993). Georeferencing of the video data made it possible to extend aerial point sampling techniques developed for 35 mm photography (Norton-Griffiths 1988) to aerial videography (Schultz *et al.* 1999a; Slaymaker *et al.* 1996).

In the mid 1990s, innovations in digital technology revolutionized video data acquisition, storage, and analysis. Two significant events, both in 1995, were the introduction of digital camcorders priced for the high-end consumer market, and the development of the IEEE 1394 High Performance Serial Bus (Jennings 1999a,b). These advances, along with improvements in multimedia capacities of personal computers, made it possible to acquire high-resolution digital video and transfer it to a computer with perfect retention of image quality, using non-professional video equipment.

This paper describes how videography acquired during four aerial surveys of the Brazilian Amazon – the Validation Overflights for Amazon Mosaics (VOAM) surveys – was processed to create a training and validation dataset for wetlands mapping with the GRFM mosaics. Although the primary goal of the video surveys was to collect ground verification for wetland cover types, we describe some further applications of the dataset, including acquisition of ground control points for image geolocation, creation of a high-resolution geocoded mosaic of a forest study area, forest biomass estimation, and rapid assessment of fire damage.

2. The VOAM surveys

Four video surveys have been carried out to date: VOAM95 and VOAM96, flown during the GRFM imaging periods for the central Amazon and limited to areas within 600 km of Manaus; and VOAM97 and VOAM99, flown during the same seasons as the earlier surveys and circumscribing the Brazilian Amazon (table 1). For VOAM95 and VOAM96, video data were acquired with a handheld Hi-8 camcorder aimed obliquely from the door or window of a light aircraft, at flight altitudes of 300–500 m. GPS location and local time were recorded by hand at approximately 1 min intervals. The camcorder display of local time was imprinted on the videotape, allowing a rough geolocation of the swath. The flight paths (figure 1a) included the floodplains of major whitewater rivers (central Solimões–Amazon, lower Purus, lower Japurá) and blackwater rivers (lower Negro, including Anavilhanas archipelago; lower Jaú); the Balbina Reservoir; interfluvial blackwater wetlands in the Jaú basin; and blocked-valley lakes (Tefé, Mamiá).

Acquired nearly simultaneously with the imaging dates of the JERS-1 mosaic scenes covering the survey flight lines (17–28 October 1995 and 24 May to 3 June 1996), the VOAM95 and VOAM96 datasets provide a permanent record of the floodplain conditions captured by the GRFM Amazon mosaics. Owing to their limited geographic coverage, however, it was clear that the 1995 and 1996 flights would have to be supplemented with additional surveys in order to support an accuracy assessment of the JAMMS basin-wide wetlands mapping. Furthermore, video data from the first two surveys has several limitations. Some footage is shaky or out of focus. Because of the analog video format, tape reproduction or frame grabbing result in significant loss of image quality, making it difficult to share the dataset with multiple users. Image scale varies across the oblique swath and as a function of aircraft height and camera angle. Although the aircraft position was recorded periodically, portions of the video swaths without distinctive landscape features cannot be reliably located. A follow-up survey was therefore flown in September and October 1997, timed to correspond as closely as possible to the JERS-1 acquisition dates in September and October 1995, but covering a greater area and using an improved camera and geolocation system to address the technical limitations of the 1995 and 1996 datasets.

For the VOAM97 survey, Brazil's National Institute for Space Research (INPE) agreed to make available their Bandeirante aerial survey plane to acquire a validation dataset for the JERS-1 wetlands mapping, for a joint INPE/UCSB analysis of the Amazon floodplain using mosaicked TM scenes (Shimabukuro *et al.* 2001) and for numerous INPE research projects in the Amazon. Approval was obtained from

Table 1. VOAM survey dates and formats. Dates of JERS-1 image acquisitions for Manaus region were 17–28 October 1995 (low-water mosaic) and 24 May–3 June 1996 (high-water mosaic).

Survey	Dates	Video format	Platform	Hours
VOAM95	18 Oct to 9 Nov 1995	One analog, oblique side-looking	Rented light aircraft	10
VOAM96	21 May to 30 May 1996	One analog, oblique side-looking	Rented light aircraft	15
VOAM97	25 Sep to 8 Oct 1997	One digital, vertical One analog, oblique rear-looking	INPE Bandeirante	35 34
VOAM99	3 Jun to 24 Jun 1999	One digital, vertical, wide-angle One digital, vertical, zoom	INPE Bandeirante	41 41
Total				176

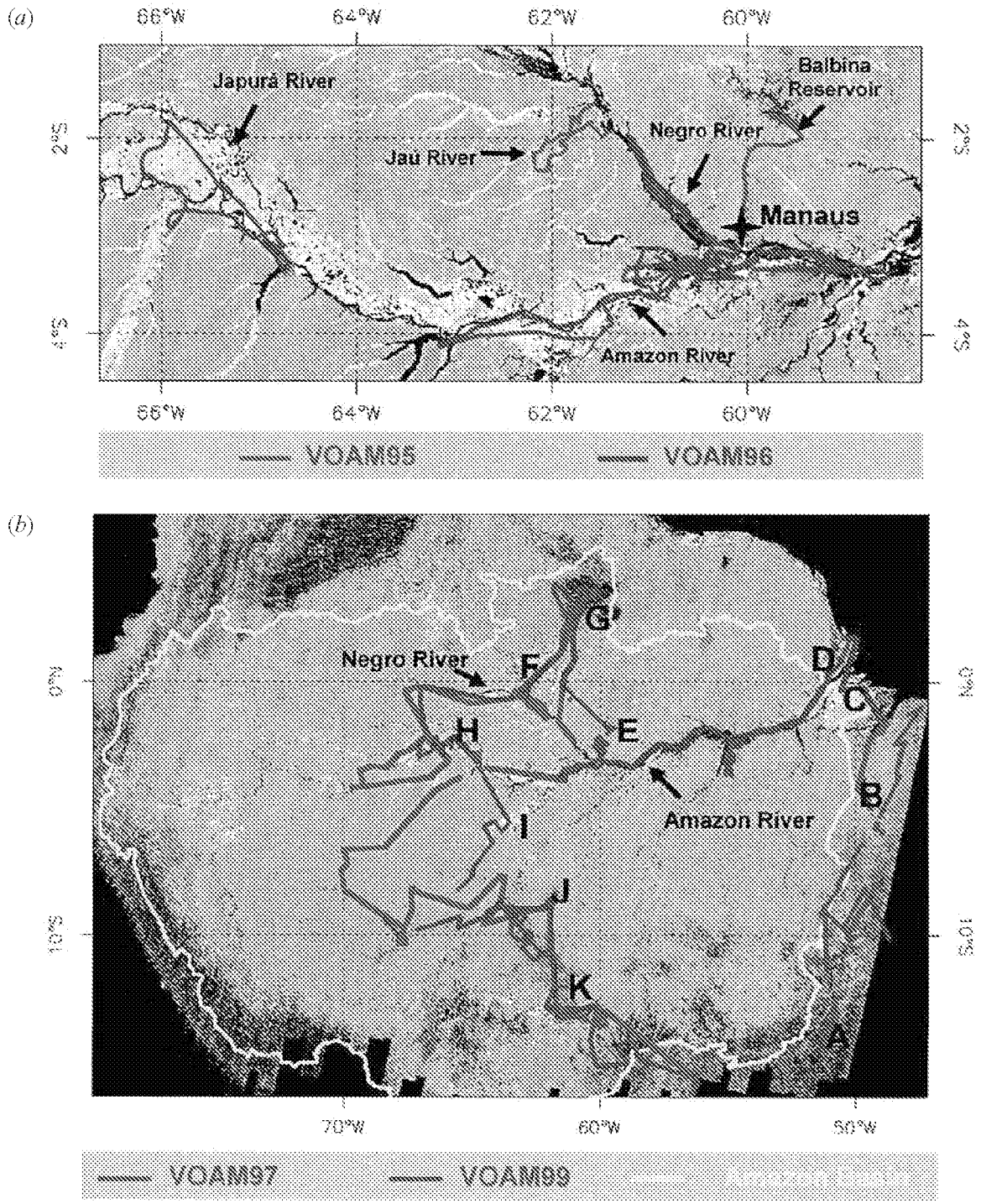


Figure 1. GRFM high-water Amazon mosaic and flight paths for VOAM surveys. (*a*) Manaus region, showing survey flight lines for 1995 (red) and 1996 (blue); (*b*) flight lines for 1997 (red) and 1999 (blue), with Amazon basin boundary shown in yellow. Major wetland and savanna features filmed on 1997 and 1999 surveys include: A, Bananal wetlands; B, Tucurui Reservoir; C, Marajó Island; D, Amapá wetlands; E, Amazon floodplain, Balbina Reservoir; F, Negro floodplain, Branco–Negro wetlands; G, Roraima savanna wetlands; H, lower Japurá floodplain; I, Purus floodplain; J, Humaitá savannas; and K, Guaporé wetlands.

Brazil's Estado-Maior das Forças Armadas for an extensive mission. The flight range of the Bandeirante made it possible to undertake a comprehensive wetlands survey of the Brazilian Amazon, covering areas remote from any airport. The flight path

(figure 1b) was designed to sample as great a variety of wetlands as possible, including the estuary, lower, middle, and upper reaches of the Amazon River within Brazil; major blackwater and whitewater tributaries spanning a wide range of latitude and longitude; campina wetlands; and major savanna regions. Although wetlands determined the general outline of the flight path, video was acquired continuously except over restricted areas, thereby sampling a broad range of undisturbed forest and savanna vegetation, pastures, various stages of forest regrowth, and urban areas.

The VOAM97 video system was assembled at the University of California Santa Barbara and the Jet Propulsion Laboratory from off-the-shelf components (figure 2). Geolocation was achieved by encoding GPS location on each video frame using a Horita GPS3 time-code generator. The time-code generator receives the aircraft's

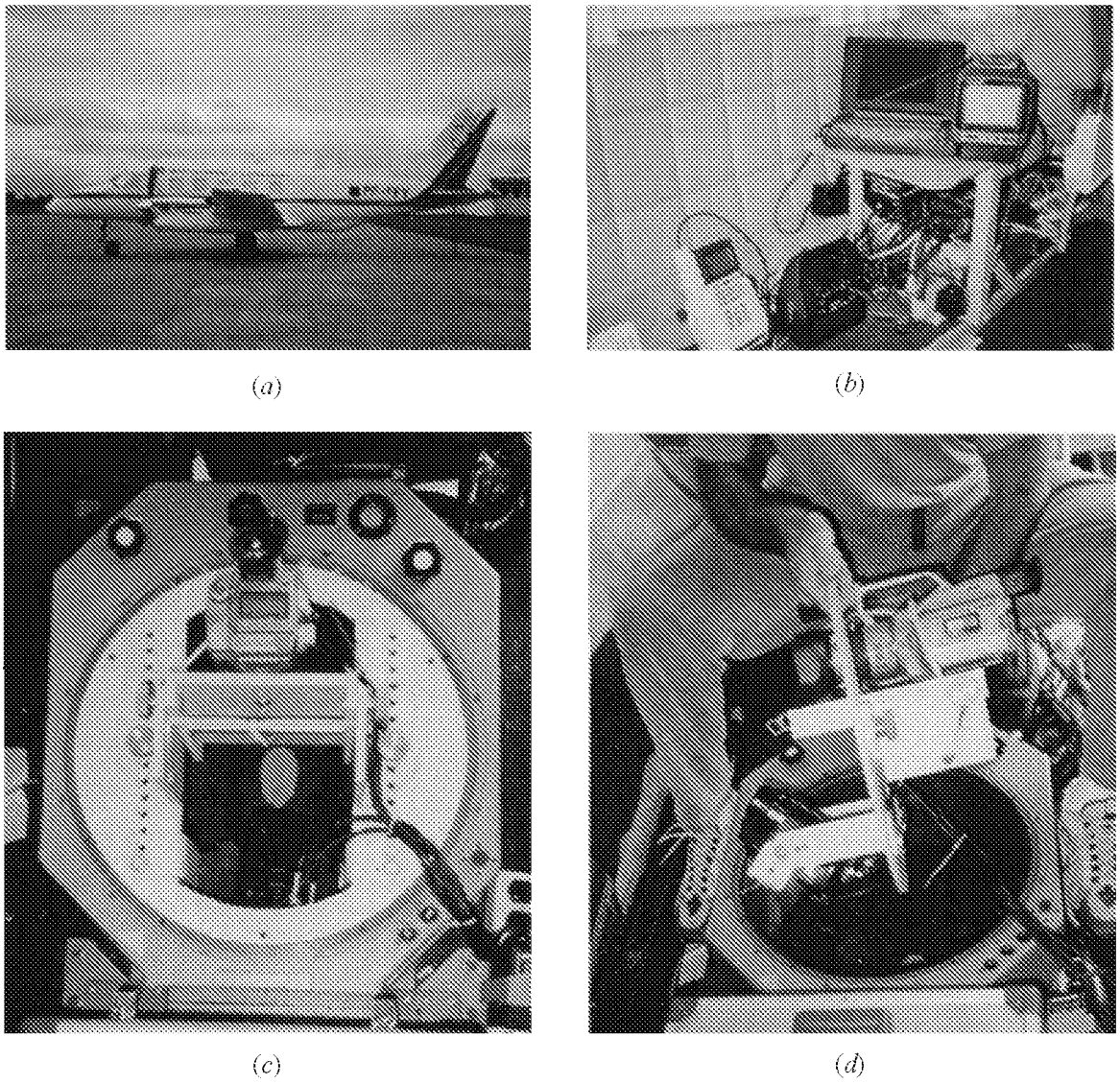


Figure 2. VOAM97 video system, as mounted in port of INPE Bandeirante survey plane (*a*). Horita GPS3 time-code generator received signal from Magellan GPS with externally mounted antenna (*b*). GPS position and time code were recorded on audio tracks of both videotapes, and in ASCII format on laptop computer. Colour monitor with switch box displayed footage from either camera, with Horita time-code display, in real time. Sony DVX-1000 digital camcorder and Sony CCD-TR500 Hi-8 camcorder were operated in tandem from camera mount custom-built by INPE (*c* and *d*).

GPS location and Greenwich Mean Time from a GPS receiver. GPS time and location are then recorded in SMPTE (Society of Motion Picture and Television Engineers) format on the audio track of the videotape. The GPS data are also recorded in an ASCII log file. Playing back the videotape with a VCR and time-code reader, the GPS time displayed on the monitor can be matched with the log file to retrieve aircraft longitude, latitude, and altitude.

The primary video dataset was acquired with a digital camcorder using DV (digital video), the most widely used consumer digital video format. Advantages of DV over analog video include improved resolution and colour accuracy, lossless duplication of videotapes and capture to disk, and smaller videotape size (Jennings 1999a). The digital camcorder was positioned in a vertical format to achieve a planimetric view. A second Hi-8 (analog) camcorder in an oblique rear-looking position was operated in tandem, to provide a broader landscape context for the vertical format swath. The Hi-8 videotape was also geolocated using the time-code generator.

The 1999 survey, again in the INPE Bandeirante survey plane, was conducted using a video acquisition system (figure 3) developed by the departments of Natural Resource Conservation and Computer Science of the University of Massachusetts (UMass) Amherst (Schultz *et al.* 1999a). A laser rangefinder (904 nm wavelength) measured aircraft height above ground or canopy, and an attitude and heading reference system (AHRS) measured aircraft tip and tilt. Aircraft position could then be projected to ground coordinates corrected for deviations from a vertical viewing angle. To improve the accuracy with which aircraft position was known, the UMass system GPS was upgraded to receive a real-time, differentially corrected GPS signal from the Omnistar commercial GPS satellite network. Both VOAM99 camcorders were digital and both were vertically oriented, one in wide-angle mode and one in zoom mode. The camcorders, laser and AHRS were mounted on an INPE-built frame. Camcorder shutter speeds were fixed at 0.001 s (wide-angle) and 0.0005 s (zoom).

Flight lines for VOAM99 (figure 1) were designed to duplicate key portions of VOAM97 flight lines and to provide coverage of intensive study sites for the Ecology component of NASA's LBA project. The LBA portion of the survey focused on human-impacted areas, including secondary forests, selectively logged forests, and pastures.

3. Processing of flight and video data

The VOAM videography can be used in its original videotape form, by viewing it on a television monitor. However, using the videotape medium it is difficult to compare different sites, make quantitative measurements, create a database of selected sites, or share data with others. All these tasks are facilitated by transferring the video data to computer disk and converting it to a non-video format. The Computer Vision Lab of the Department of Computer Science at UMass has developed a suite of software tools to automate frame extraction, mosaicking, and geocoding of the video data and maximize its usefulness for environmental monitoring applications (Schultz *et al.* 1999a, b, Zhu *et al.* 1999). This section discusses transfer of video data from tape to computer disk and processing with the UMass programs, using the VOAM99 dataset as an example. A diagram of the processing flow is shown in figure 4.

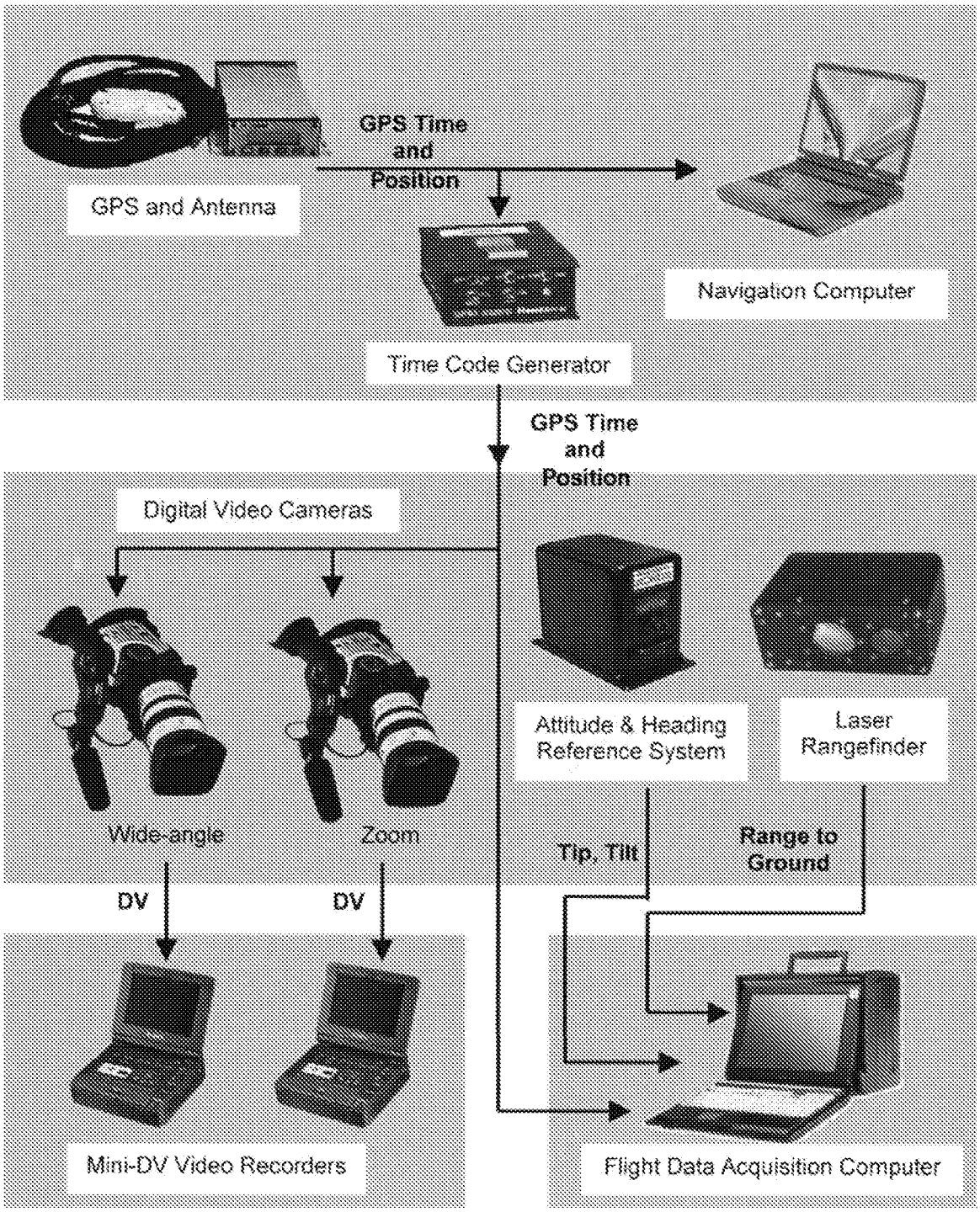


Figure 3. Digital video acquisition system for VOAM99 survey. OmniSTAR real-time, differential GPS signal received by Trimble 132 GPS using externally mounted L-band receiver was converted to SMPTE format by Horita GPS3 time-code generator, for encoding on audio tracks of digital videotapes. GPS data were also used for continuous updating of flight track on navigation computer. Wide-angle and zoom format videography acquired simultaneously by two Canon XL-1 digital camcorders was recorded in mini-DV format and viewed using Sony GV-D900 Video Walkman VCRs. Data streams from Watson Industries BA303 AHRS (attitude and heading reference system), Laser Atlanta Advantage laser rangefinder, and GPS were time-stamped and recorded by flight data acquisition computer for post-flight processing.

3.1. *Flight data processing*

The flight data acquisition computer recorded data streams from the laser altimeter (sampling at 238 Hz), the AHRS (11.8 Hz), and the GPS time-code generator (30 Hz, determined by the video acquisition rate of 30 frames per second). Because of the different sampling rates, the streams were recorded separately, with a system clock time stamp for each reading. Post-processing was therefore required in order to synchronize the data streams and to resample all data fields to the sampling rate of the laser. The AHRS and GPS location fields were interpolated between readings. Since video rate is fixed at 30 frames per second, video time codes (hour:minute:second:frame) were repeated rather than interpolated. Each video frame is thus associated with multiple laser, AHRS, and GPS location fields. Using AHRS tip and tilt readings and the laser measurement of distance from aircraft to ground, aircraft position readings from the GPS were corrected for the offset between the GPS antenna and camera position and projected to give the ground positions of the laser hits. The WGS84 datum was used for all geographical coordinates.

Absolute elevations of laser hits were calculated by subtracting the vertical distance between aircraft and ground from GPS-determined aircraft elevation. For portions of the survey flown at altitudes beyond the range of the laser (350–400 m above ground), an alternate flight data processing option allows the user to enter an estimated ground elevation, from which approximate aircraft-to-ground distance can be calculated and used to estimate ground position. Ground elevation can be estimated from a map or from occasional laser returns from highly reflective objects, detectable at altitudes beyond the normal laser range.

3.2. *Capture of video data to computer disk*

The date and time of the video footage for a particular site are retrieved using flight-line vector and attribute files generated from flight logs. Flight lines are overlaid on a georeferenced satellite image and time codes are selected by clicking on points along the flight-line and viewing their attributes. The videotape is viewed using a digital VCR (video cassette recorder) with video output routed through a time-code reader, which displays on the television monitor the GPS time code that was encoded during the flight on the audio track of the videotape. The user can then advance to the correct portion of the videotape to capture footage of the target site.

For video capture, the VCR is controlled from the monitor of a computer equipped with digital video capture card and software, connected to the VCR with an IEEE 1394 connector. The transfer to computer does not entail decompression or alteration of the original format; the DV format is retained on disk, and the transfer is lossless. The DV file may be 'wrapped' by the capture software in the format of a multimedia architecture; for example, the MotoDV capture card used to process the VOAM datasets wraps DV clips so as to be readable with the QuickTime multimedia utilities. DV files require large amounts of disk space—about 13 Gb for 1 h of video.

3.3. *Frame extraction*

Various kinds of multimedia software can be used to extract the frames from a video clip and create image files for each frame, but they must be modified in order to retain the GPS information from the audio track. The UMass frame extraction program, built on QuickTime file-handling utilities, automatically names each image file using the acquisition date and GPS time code and frame number so that the

image filenames alone give an exact cross-reference to the flight log data. The frame extraction program creates 24-bit BMP files with dimensions of 720×480. Each file is approximately 1 Mb; a 30 s clip requires about 900 Mb of disk space for these intermediate files.

3.4. *Mosaicking and geocoding*

Once converted to image format, the individual frames extracted from the video clip can be mosaicked into a continuous strip with inexpensive commercial mosaicking programs that use frame matching to build the mosaic. Drawbacks to this 'free' mosaic method are that small errors accumulate during the mosaicking, resulting in large positioning errors towards the end of the strip; changes in scale owing to topographic relief are not reflected in image scale; and geocoding of strips (post-mosaicking, using control points derived from flight logs) is labour-intensive. However, the mosaicking algorithm in the UMass software (Zhu *et al.* 1999) addresses these problems by using the flight log data to create a globally correct model of the motion track, constraining the propagation of errors. A 'two-track' method is employed as follows:

1. Initial registration of frames. Matches within adaptive search windows in current and reference frames are evaluated using a combination of correlation (normalized absolute difference) and texture (normalized average magnitude of the gradient image). Inter-frame motion parameters are iteratively estimated using a weighted least mean squares method.
2. Track generation and refinement. An 'estimated' track is generated from the initial registration, and an 'expected' track from the flight log data. The estimated track will yield a mosaic that is seamless but without geographical reference, while the expected track will yield a mosaic that is globally consistent with geographical data from the flight logs, but has discontinuities between frames. Comparison between the two sets of inter-frame motion parameters is then used to refine and update the two tracks.
3. 'Two-track' mosaic composition. The geocorrected mosaic is built by applying a line-by-line transformation to the seamless mosaic produced from the refined estimated track, using the refined expected track to correct accumulated geographical error frame by frame. The resulting mosaic is globally correct geographically, but maintains the precise seamless stitching of successive frames achieved by the initial frame registration.

The UMass program used to mosaic VOAM99 video clips outputs the mosaic as a 24-bit file in GeoTiff format (Ritter and Ruth 1997). An automatically generated GeoTiff header gives pixel dimensions in metres, and UTM coordinates of the upper left corner of the mosaic.

3.5. *Processing of earlier VOAM datasets*

Footage from the 1997 survey was processed in a similar manner to VOAM99 data. Since the VOAM97 system did not include a laser profiler or AHRS, ground positions of frame centres were assumed to be equal to aircraft positions, and scale differences due to variations in terrain elevation were not incorporated in the mosaicking. The oblique footage from the 1995, 1996, and 1997 surveys cannot be mosaicked into geocoded image files as can the vertical format data. However, clips from these datasets can be transferred to computer disk by first copying them from

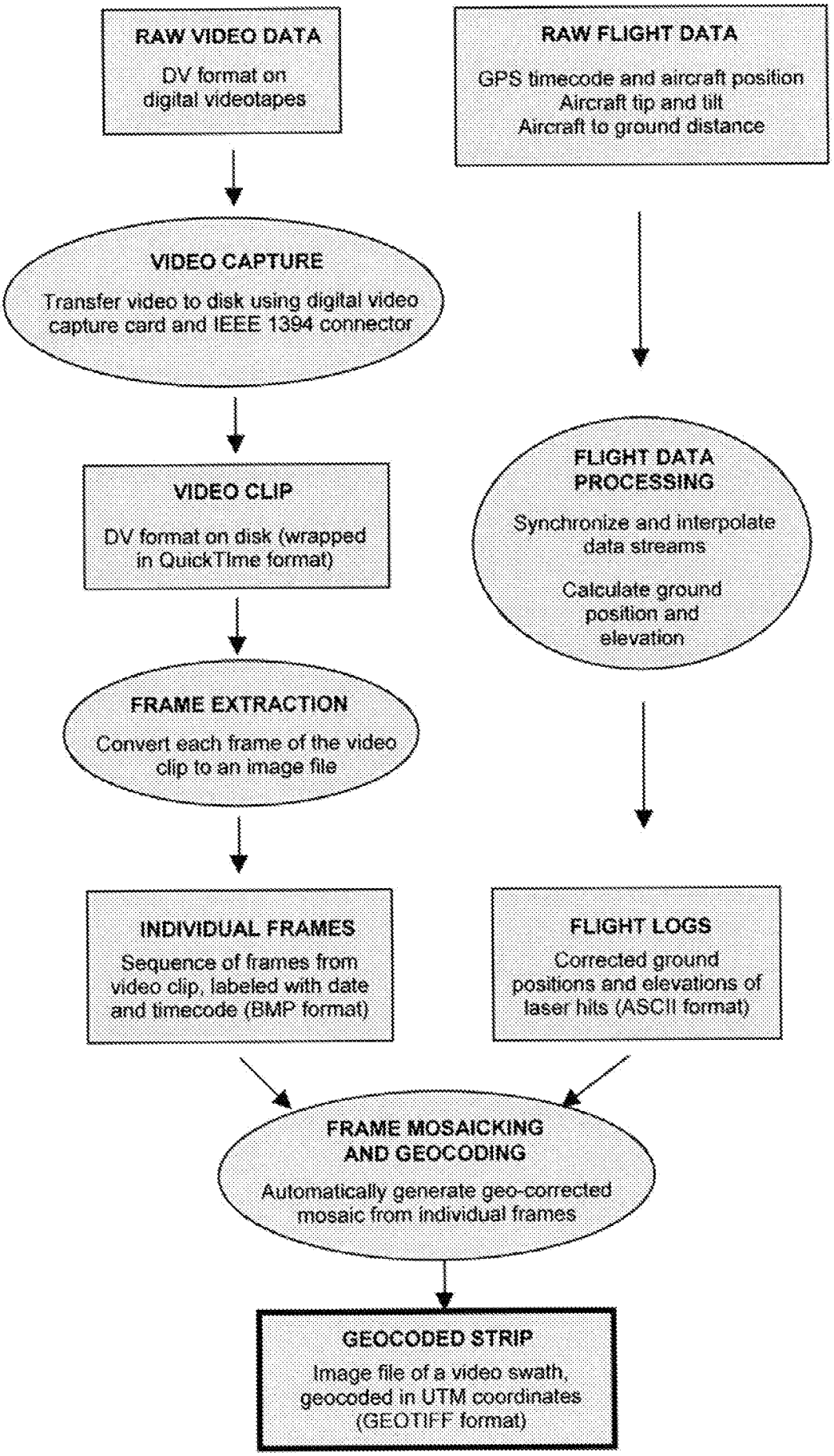


Figure 4. Processing flow for digital video mosaicking and geocoding.

their original Hi-8 format to DV format videotapes. Frames can then be extracted, and for some footage short mosaics can be constructed. Frames or mosaics can be approximately geolocated by cross-referencing with flight log data.

4. Characteristics of the video datasets

4.1. *Timing of surveys relative to flooding seasonality*

The timing of the VOAM surveys and JERS-1 acquisitions relative to Amazon River stage is shown in figure 5 for Manacapuru gauge, located 60 km west of Manaus. JERS-1 high-water mosaic images for the central Amazon were acquired at river stages 30–60 cm higher than the long-term mean high-water stage, while low-water mosaic images were acquired at river levels from 2.2 m to 2.6 m below mean low-water stage. The recurrence interval for river stage as low as that captured on the JERS-1 low-water mosaic is approximately 10 years. At such stage levels, the beds of shallow floodplain lakes and depressions usually inundated year-round are exposed, as are sandbars and other depositional features in river channels (figure 6a). Interannual variability in Amazon River discharge appears to be coupled to El Niño–Southern Oscillation cycles (Richey *et al.* 1989).

Because the VOAM95 and VOAM96 surveys were carried out during the JERS-1 imaging periods, floodplain inundation conditions recorded during those surveys correspond to those on the mosaics. However, stage levels were about 1 m higher during the VOAM99 survey than during the JERS-1 high-water imaging, and 1–2 m higher for the VOAM97 survey than for the JERS-1 low-water imaging. Thus, some areas that were inundated during the VOAM97 survey, particularly in shallow lakes, were exposed on the low-water mosaic, and some higher areas on levees were dry on the high-water mosaic but flooded during the VOAM99 survey. These discrepancies may introduce errors during accuracy assessment. For the topographically highest and lowest parts of the floodplain, the VOAM95 and VOAM96 surveys, though of lower resolution and poorly geolocated, may provide better training and

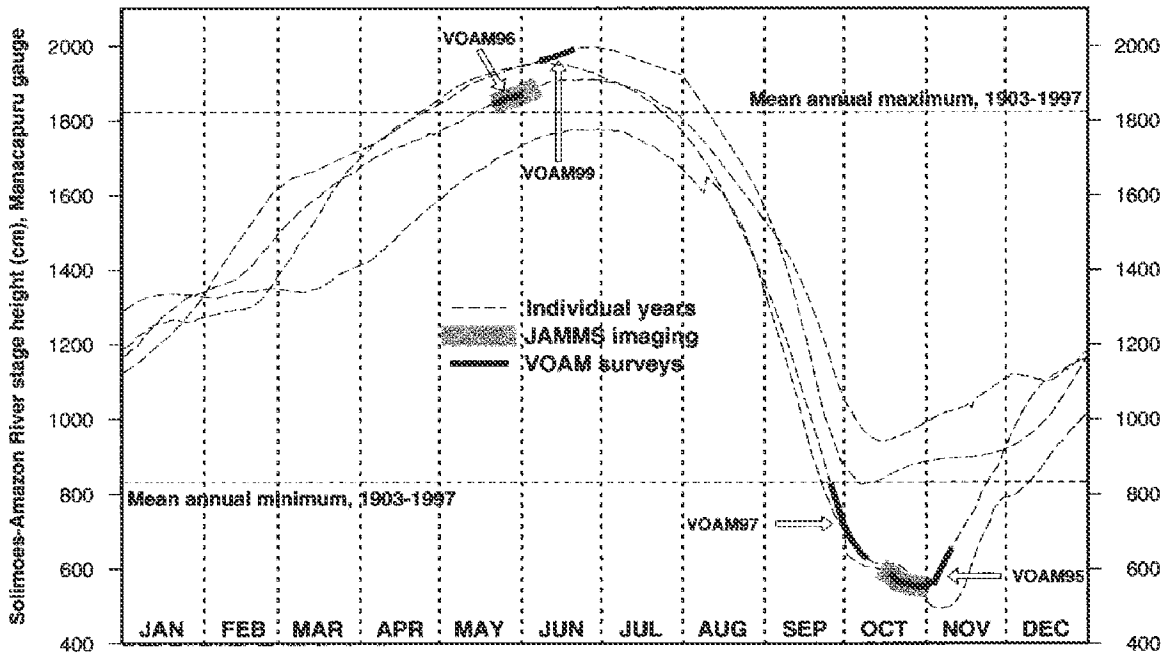


Figure 5. Timing of VOAM surveys relative to Amazon River stage at Manacapuru gauge, approx. 60 km west of Manaus.

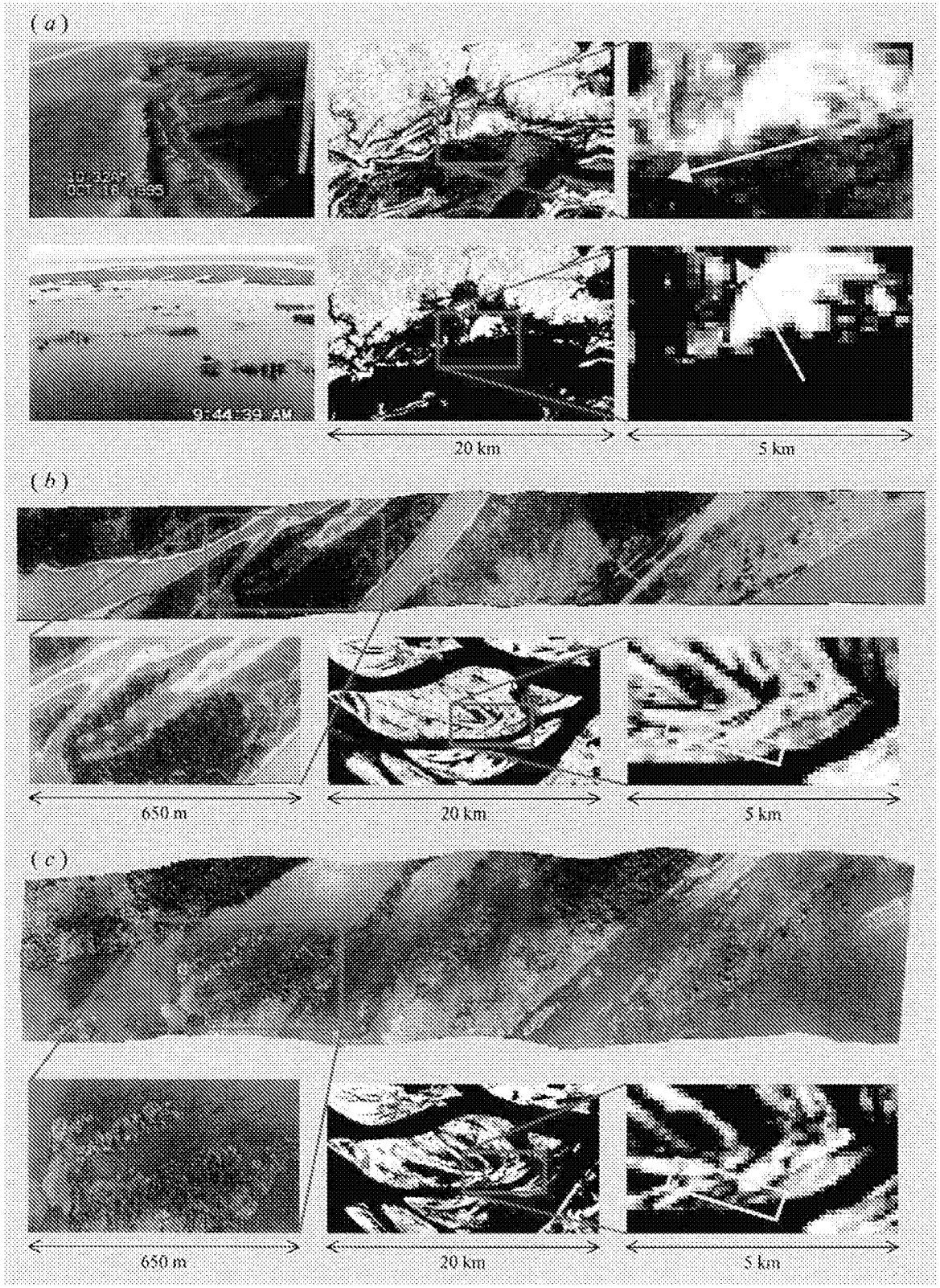


Figure 6. (*a*) Captured analog video frames from VOAM95 survey for Lago Cabaliana (left) and Anavilhanas Archipelago of Rio Negro (right). Cabaliana view is west-looking along northern shore of lake; left portion of frame shows grasses growing on lake bed, exposed only at below-average low-water levels. VOAM97 (low-water, *b*) and VOAM99 (high-water, *c*) digital video mosaics are for Amazon floodplain site near Parintins.

validation datasets. However, given the 13 m difference in Amazon River stage between the low- and high-water mosaics, the VOAM97 and VOAM99 surveys adequately represent water levels on the JERS-1 mosaics for the central Amazon.

The contrast between low- and high-water conditions for a typical Amazon floodplain site is seen in the two mosaicked swaths shown in figure 6. For this site near Parintins, várzea forest stands and individual larger trees appear the same at low water (6*b*) and high water (6*c*), while shorter trees and shrubs are submerged at high water. Areas with bare soil and grass at low water have either open water or floating macrophyte cover at high water.

4.2. *Resolution and swath width*

Flying height for the VOAM99 survey ranged from 100 to 900 m, with a median of 350 m. For segments where laser data was required, flying height was selected to maximize swath width while remaining within the range of the laser instrument and below the cloud ceiling. At the median flying height, swath width was about 550 m (wide-angle) and 55 m (zoom), with pixel dimensions of about 0.75 m (wide) and 7.5 cm (zoom). Resolution of the VOAM99 video data was in general equivalent to pixel dimension, with the exception of occasional loss of focus over smooth water surfaces. On the high-resolution zoom footage, features such as individual branches, leaves, and grass stems are distinguishable.

Swath width and pixel dimension were similar for the VOAM97 vertical format data, but resolution was lower for three reasons. First, focusing problems occurred intermittently during the 1997 survey; during these episodes the camera went slightly out of focus for a few frames before returning to focus. Secondly, the digital camcorder used for the 1997 survey used interlaced video, in which odd and even lines of a frame are recorded at different moments; this can cause blurring when the camera is moving relative to the target, as in aerial surveys. The progressive scan video used in the VOAM99 camcorders does not use interlacing. Thirdly, the high levels of atmospheric haze and smoke typical of the late dry season reduced the quality of the 1997 data, particularly in the oblique format.

For the oblique VOAM95 and VOAM96 datasets, resolution and swath width were variable, depending on camera angle and zoom factor.

4.3. *Geolocation accuracy of VOAM97 and VOAM99 data*

The absolute geolocation accuracy of mosaics generated from VOAM97 and VOAM99 data is affected by errors from several sources.

1. *Aircraft GPS location.* For the real-time, differentially corrected GPS signal received once per second during the VOAM99 survey, expected probabilities of horizontal error are 67–73% that location error will be less than 0.5 m, 95–97% that it will be less than 1.0 m, and 99% that it will be less than 1.5 m. The range of vertical error is 2–2.5 times as great, resulting in a 67–73% probability of vertical error less than 1.25 m, 95–97% probability of error less than 2.5 m, and 99% probability of error less than 3.75 m. For the VOAM97 survey, which was based on non-differentially corrected GPS readings every 2 s, horizontal error probabilities are estimated at 95% for errors < 100 m.
2. *Frame centre location, pre-mosaicking.* For VOAM99, calculation of the horizontal offset between aircraft location and frame centre location is affected by errors in GPS-based vertical position of the aircraft, in range to ground

(measured by laser rangefinder), and in aircraft tip and tilt (measured by AHRS). At an aircraft height above ground of 350 m and aircraft tip/tilt of 5°, and assuming instrument errors of ±2.5 m from the GPS, ±15 cm from the laser rangefinder and ±0.3° from the AHRS, the resulting error in offset calculation would be 3 m, 2.5 m of which is owing to error in tip/tilt measurement. Adding this to the 1.0 m horizontal error achieved 95–97% of the time for aircraft position yields an error of up to 4 m in horizontal positions of frame centres. Air turbulence would add further error to frame centre location. The 1997 survey did not include AHRS or laser measurements. Again assuming an aircraft height above ground of 350 m and aircraft tip/tilt of 5°, a horizontal error of ±43 m would result from the lack of correction owing to tip/tilt.

3. *Pixel location, post-mosaicking.* The amount of error introduced during mosaicking is estimated to be on the order of two or three pixels, or about 1–3 m for wide-angle mosaics at typical flight altitudes, for the centre track of the mosaic. Relative and absolute geolocation errors can be expected to increase along the margins of mosaics, owing to lens distortions and to terrain height effects where terrain height varies across the video swath. Optical distortion models for the lenses used in the two surveys are not currently available.

Because of the number of interacting factors affecting geolocation accuracy, and the large amount of data collected, a comprehensive assessment of geolocation error is not feasible for the VOAM97 and VOAM99 datasets. Comparison with field GPS measurements for a small number of sites indicates that absolute errors are in the range of 5–10 m (1999) and 50–150 m (1997) along the centre third of mosaicked swaths.

4.4. *VOAM99 laser data*

The laser altimeter recorded the range to the last return from each pulse. At the sampling rate of 238 Hz, this resulted in about eight laser hits per video frame. The power of the laser was limited to meet eye safety standards when flown at low altitudes. The percentage of 'zero returns'—pulses for which no return was detectable—increased with aircraft height above ground as shown in figure 7. Over terrain without standing water, at aircraft heights above ground of less than 200 m, less than 10% of returns were below the detection threshold; zero returns increased to 17% at 300 m and 37% at 400 m. Segments with more than 20–30% zero returns are of limited use for analysis of vegetation heights, but can be used for mosaic geocoding. Returns from water surfaces varied with water surface roughness and presence of sunglint; reflectivity from smooth water surfaces without sunglint was below detection thresholds. At a range of 350 m, the beamwidth of the laser pulse is approximately 1 m, allowing it to penetrate small gaps in vegetation canopies and record range to ground or understorey.

Figure 8 shows a laser profile along a 37 km flight-line segment in the southern portion of the Roraima–Rupununi savanna region. Beginning on the north bank of the Cauamé River, the transect extends south-west across the river floodplain, through the outskirts of the city of Boa Vista, and across several streams that drain south-east to the Branco River. This area is part of an extensive, largely undissected pediplain, having a low-density, non-organized drainage pattern with closed lakes (Schaefer and Dalrymple 1995). The flatness of the terrain can be seen in the laser

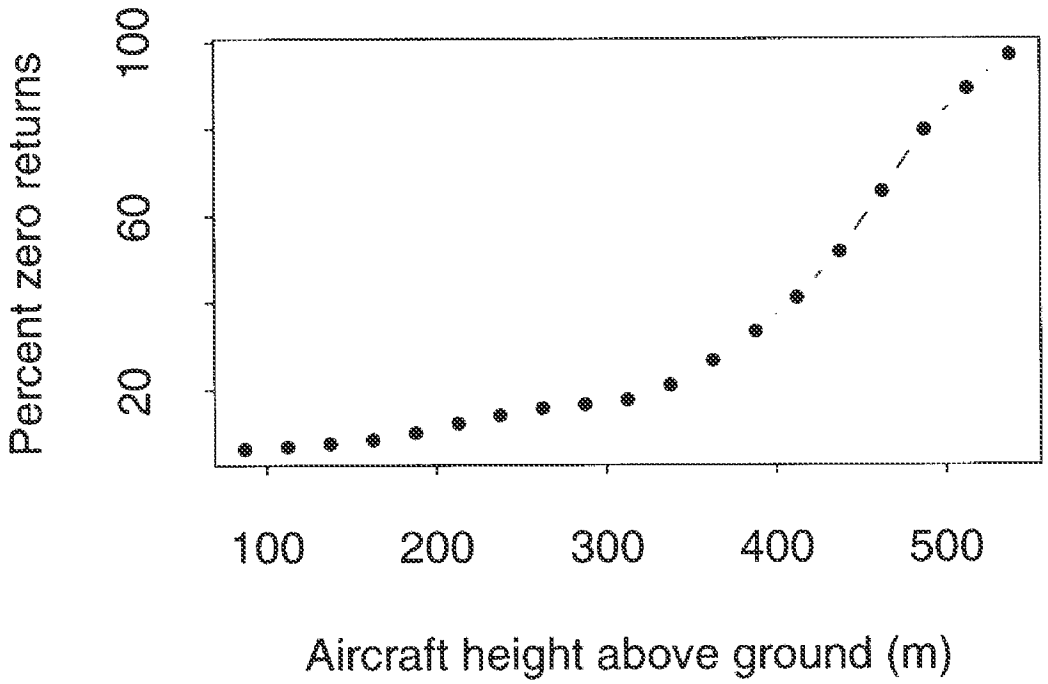


Figure 7. Per cent of pulses from VOAM99 laser rangefinder for which no return was detectable ('zero returns') as a function of flying height above ground.

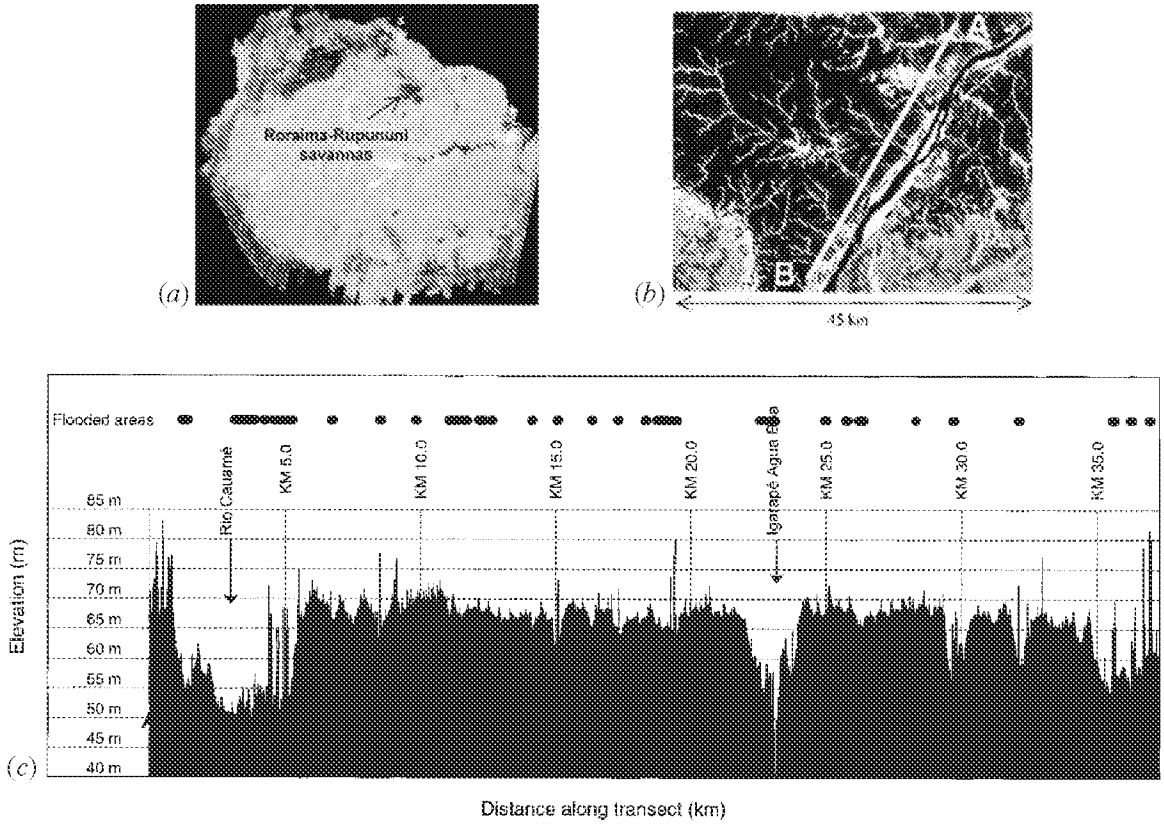


Figure 8. Laser profile of 37 km transect across Roraima savannas (location shown in *a*). Flooded waterways with woody vegetation, mainly *Mauritia flexuosa* palms, are seen as bright stream network on JERS-1 high-water mosaic (*b*), and as flooded depressions with laser hits from treetops on laser profile (*c*). Points plotted across top of profile show areas where flooding is visible on videography.

profile: ground elevations outside of drainages decrease from 67 m in the north-east to 65 m in the south-west.

Portions of the transect for which standing water is visible on the videography are indicated along the top of figure 8. The upland surface is about 15 m higher in elevation than the river channel, and more than 10 m above peak flow level for the river. The presence of standing water in upland depressions and drainageways indicates that direct rainfall input and local runoff are the primary inputs to surface flooding, and that seasonality of flooding in areas away from the river floodplain cannot be directly predicted from river stage. In addition to topography, the laser profile gives information about vegetation heights, showing shrub heights of 1–5 m and tree heights of 10–15 m (palms) and 10–20 m (other broadleaved) along the transect.

5. Use of digital videography to train and validate wetlands masking and classification

5.1. *Creation of the JAMMS wetlands mask*

The JAMMS wetlands mapping is done in two phases: (1) a wetlands mask is created that excludes all non-wetlands features; (2) the areas within the wetlands mask are classified into cover types on both the low- and high-water mosaics. Wetlands are defined as vegetated or non-vegetated areas that are seasonally or permanently inundated. The masking and classification steps are being applied to mosaic tiles measuring 6° of longitude by 4° of latitude (7200×4800 pixels at the JERS-1 mosaic resolution); the JAMMS study area comprises 24 of these tiles. Here we describe use of the VOAM99 data for assessing the thematic accuracy of the wetlands mask for the tile containing the city of Manaus.

The method for creating a wetlands mask from the co-registered low- and high-water JERS-1 mosaics is described by Barbosa *et al.* (Barbosa *et al.* 2000). Discrimination between wetland and non-wetland features based on the radar backscattering coefficient (σ°) of JERS-1 mosaic pixels does not yield good results, since many wetland and non-wetland features display similar L-HH backscattering statistics. Examples of wetland/non-wetland pairs with similar L-HH returns include flooded palms and urban areas (both with σ° ranging from -3 to -2 dB); flooded várzea forest and facing hillslopes oriented normal to the JERS-1 look direction (-6.5 to -4 dB); and aquatic macrophytes and pastures (-14 to -9.5 dB). Such pairs of features cannot be well separated by pixel-based classifiers using σ°, texture of σ°, or ratio of low- to high-water σ°; however, they are readily distinguished by a human interpreter based on spatial pattern and context. A hybrid machine-interpreted and human-interpreted method was therefore employed to generate the wetlands mask.

The co-registered JERS-1 mosaics were segmented into homogeneous polygons using a region-growing algorithm, and the resulting polygons were assigned as wetland or non-wetland after an unsupervised clustering using polygon backscattering statistics. The image segmentation and polygon-based classification routines used are part of INPE's SPRING software package (Camara *et al.* 1996). The automatically generated and clustered polygons were edited by a human interpreter. With this polygon-based approach, the interpreter can incorporate spatial pattern and context into editing decisions, which would not be practical with a pixel-based classification.

Classification of a polygon as a wetland was thus based initially on backscattering values indicative of flooding on either high-or low-water mosaics, then modified based on location or spatial pattern indicative of a wetland or non-wetland, as

judged by a human interpreter. Cases where assignment of a polygon as wetland or non-wetland was in doubt were resolved in most cases by comparison of the mosaics with TM imagery or maps. In some instances, VOAM videography was viewed in order to make an assignment; the video segments used for this purpose were excluded from the dataset used for validation of the wetlands mask. The resulting wetlands mask for the Manaus tile is shown in figure 9; 31% of the tile area is covered by wetlands. The minimum mapping unit (equivalent to the minimum polygon size specified during the image segmentation) is 50 pixels, or approximately 0.43 km^2.

5.2. *Creation of the validation dataset*

Sample locations for assessing the thematic accuracy of the wetlands mask were selected by a random sampling of flight times along the VOAM99 survey tracks, stratified by class (i.e. wetland or non-wetland) and using a minimum time sampling unit of 1 s. Each sample consists of a 100×100 m area centred on the UTM coordinates corresponding to the randomly selected flight time. Two constraints were placed on the sample selection: samples could not be located within 10 flight seconds of one another, to avoid overlapping samples; and a maximum of two samples was allowed within each cell of a 5×5 km grid superimposed on the tile, to avoid oversampling areas with multiple flight lines.

The sample size required to achieve a confidence level of 95% and a precision of 2.5% for estimation of error matrix components was calculated to be 1286, based on the multinomial distribution (Congalton and Green 1999, Tortora 1978). Class proportions of 80% non-wetland and 20% wetland were assumed for this calculation, estimated from results of a preliminary wetlands masking of six mosaic tiles. VOAM99 flight lines traversed 13 of the 24 mosaic tiles. Of the 1286 randomly selected flight times, 145 fell within the Manaus tile (figure 9), which had greater

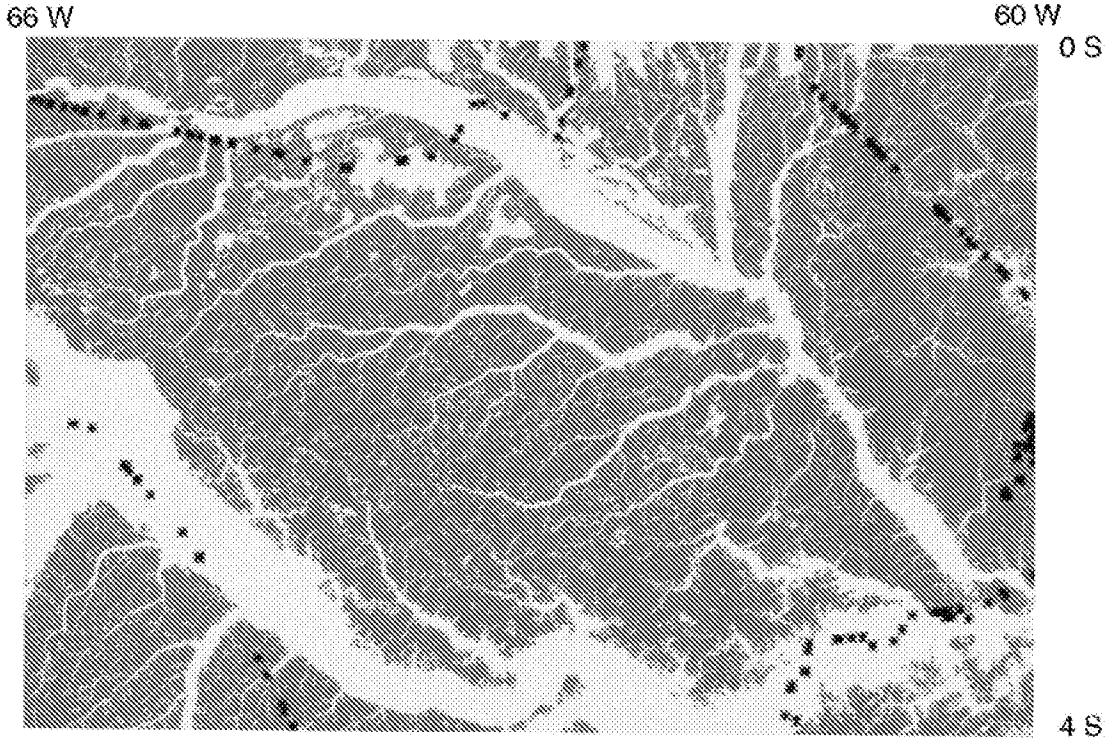


Figure 9. Locations of validation sample points (101 non-wetland, 44 wetland) on Manaus wetlands mask.

than average flight coverage. This represented 0.5% of the sample pool of 30,051 flight times within the tile, after excluding survey segments referenced during mask editing.

Ten-second wide-angle and zoom video clips centred on the randomly selected sample times were saved to disk and the samples were labelled as wetland or non-wetland, with the exception of samples for which (a) the aircraft was banking at a steep angle during a turn (about 4% of the random samples); (b) the ground was obscured by clouds, smoke, or closure of the belly port during rainfall (3% of samples); (c) the sample could not be identified as either wetland or non-wetland (7% of samples); or (d) the sample area fell directly on a wetland/non-wetland boundary (less than 1% of samples). Labelling was based solely on the video record. Samples were labelled as wetland if they were flooded during the VOAM99 survey, or if they had vegetation types or geomorphologic traits diagnostic of Amazonian wetlands, even though they were not flooded during the video survey.

Areas flooded during the survey were readily detected in most cases, either by sunglint from water surfaces (seen through canopy gaps in flooded forests), or by reflections of sky, cloud or airplane in the water surface. Flooding was more difficult to detect on overcast days. In some campina wetlands (blackwater wetlands on white sand), patches of bare, dry sand were difficult to distinguish from sunglint or cloud reflections. Some flight-line orientation and sun angle combinations were difficult to interpret, especially at higher flight altitudes.

Labelling a sample as non-wetland was more time-consuming than labelling a wetland sample, since it required confirmation of the absence of flooding, and since upland forest or shrub canopies may superficially resemble wetland canopies. In many cases, it was necessary to view the zoom as well as the wide-angle footage, to verify lack of flooding by close viewing of ground features in canopy gaps. For footage shot at high flight altitudes and low sun angles, ground features were likely to be obscured by shadow, making it difficult or impossible to rule out the presence of flooding.

Since the JERS-1 high-water mosaic and VOAM99 survey were timed to peak high water for the central Amazon, it is not surprising that for the Manaus tile, all samples labelled as wetland were identifiable on the video by the presence of flooding. In parts of the Brazilian Amazon with a different seasonality of inundation, wetlands may not have been flooded during the VOAM99 survey. In those cases, distinctive vegetation, or geomorphologic features such as meander scars and oxbow lakes, will aid in labelling wetlands samples that were not flooded during the survey.

For each of the 145 sample points, the validation dataset record consists of the wide-angle and zoom video clips in QuickTime format, mosaicked wide-angle and zoom strips in GeoTiff format, and flight log data in ASCII format. Figure 10 shows the geocoded wide-angle mosaic and an enlarged portion of the zoom mosaic for a sample located in a campina wetland near the confluence of the Negro and Unini rivers, and inundated during the video survey. The video clips, which require large amounts of disk space, are retained as part of the sample record because flashes of sunglint and reflections in the water surface are much more obvious in the dynamic video format than on the static mosaics.

5.3. *Assessment of thematic accuracy of the wetlands mask*

The error matrix for the Manaus tile validation is shown in table 2. For each of the 145 samples, the videography-based class assignment (wetland or non-wetland)

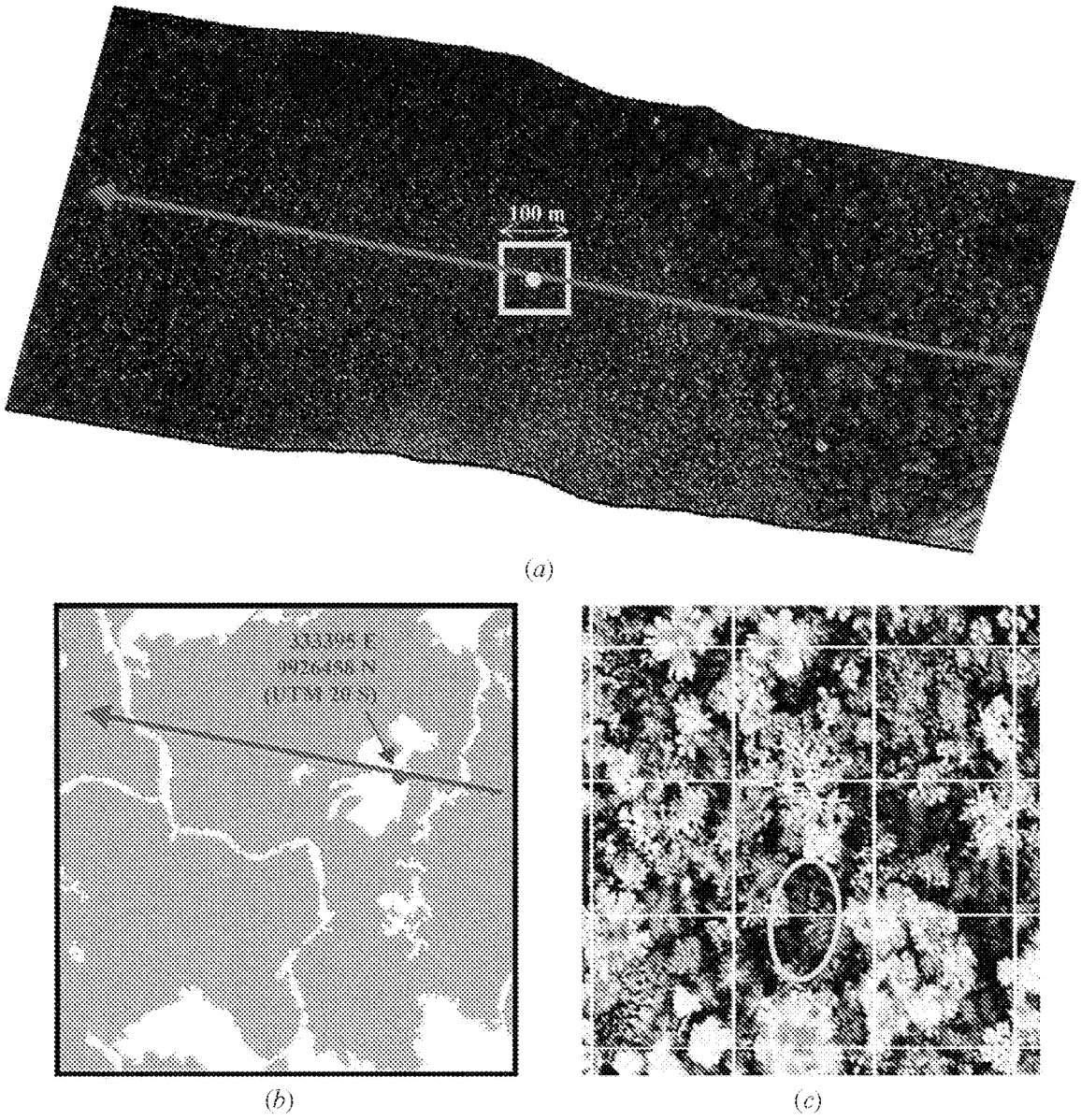


Figure 10. Validation sample for JAMMS wetlands mask, Manaus tile. Interfluvial wetland on white sand near confluence of Negro and Unini rivers, 18 June 1999. The 1 ha sample is centred on a randomly selected point along the flight line. (*a*) Mosaic of 10 s VOAM99 flight segment (wide-angle), showing flight line (red) and sample area (yellow). Wetland/upland boundary extends diagonally from lower right to top centre of mosaic. The smooth texture of the campina wetland contrasts with the rough texture of the upland forest canopy. (*b*) Location of sample on VOAM99 flight line, plotted on wetlands mask (white = wetland, grey = non-wetland). (*c*) Portion of zoom mosaic, with blackwater inundation visible through mixed shrub and palm canopy.

was compared with the classification of that 1 ha area on the wetlands mask. Four of the 101 non-wetland samples were misclassified on the mask as wetland, and one of the 44 wetland samples was misclassified as non-wetland, an overall accuracy of 96.5%. Producer's accuracies for non-wetland and wetland were 96% and 98%, respectively; producer's accuracy estimates the probability that a given reference sample has been correctly mapped on the wetlands mask. User's accuracies, which give the probability that a given mask sample has been correctly mapped, were 99%

Table 2. Error matrix, producer's accuracy, and user's accuracy for Manaus wetlands mask, based on reference data derived from VOAM99 digital videography.

		Reference data		
		Non-wetland	Wetland	Total
Wetlands mask	Non-wetland	97	1	98
	Wetland	4	43	47
	Total	101	44	145

Producer's accuracy		User's accuracy	
Reference	Per cent	Mask	Per cent
Non-wetland	96	Non-wetland	99
Wetland	98	Wetland	92

Overall accuracy = 140/145 = 96.5%.

for non-wetland and 92% for wetland. All misclassified samples were either located within one to two pixels of a polygon boundary or were located on features smaller than the minimum mapping unit used for the wetlands mask, e.g. small non-wetland islands or peninsulas surrounded by wetland. Thirty per cent of the 145 reference samples for the Manaus tile were wetland. This is a higher wetland cover than the 20% estimate for the study area as a whole because the Manaus tile has a higher than average concentration of large river floodplains, and is nearly equal to the 31% wetland cover estimated by the mask for the Manaus tile.

Although the video-derived validation samples were randomly selected from VOAM99 flight times, the flight lines themselves were purposefully located to maximize wetland and LBA site coverage while minimizing total flight distance. The validation dataset therefore must be considered a probability sample of the subregion covered by the videography, but a nonprobability sample of the study region as a whole (Stehman and Czaplewski 1998). Given the large area and small number of airports in the study region and the desired focus on wetlands, systematic or random location of the flight lines was not practical. The area covered by the survey can be viewed as representative of the study area in that the flight paths were designed to sample the full range of wetland types of the Brazilian Amazon, but cannot be proven to be representative in a statistical sense. Further videographic surveys are being planned to obtain coverage over large Amazonian wetlands for which the VOAM survey data (confined to Brazil) may not be representative, such as the Marañon–Ucayali palm swamps of Peru and the Bolivian Llanos de Moxos.

5.4. *Training of classification algorithms for JAMMS wetlands vegetation and inundation mapping*

The second phase of the JAMMS wetlands mapping—classification of the areas within the validated wetlands mask into vegetative–hydrologic classes—is ongoing. After a second image segmentation of the co-registered mosaics (excluding the non-wetland portions), statistics are extracted for the segmentation-based polygons. A rules-based classifier is then derived based on a training set of segmentation-based polygons, and applied to the full set of wetlands polygons. The vegetative–hydrologic

classes are based on vegetation structure and inundation state (open water, bare ground, non-flooded or flooded herbaceous, non-flooded or flooded woodland, non-flooded or flooded forest). For some of the wetlands classes, JERS-1 backscattering statistics are not unique on a single date, but can be distinguished using the two-date signature. For example, the VOAM99 flight segment shown in figure 11 traverses three training polygons on the segmented two-date mosaics. The median backscattering coefficients (σ°) for polygons A, B, and C respectively are -7.3 dB, -6.0 dB, and -7.3 dB at low water, and -5.6 dB, -7.3 dB, and -7.9 dB at high water.

The videographic record aids in interpreting these signatures and correctly training a classifier. The wide-angle and zoom frames show that polygon A is closed-canopy várzea forest, for which σ° increases from -7.3 to -5.6 dB when the forest is flooded at high water. This increase of 1.7 dB is expected because the smooth water surface in the flooded state acts as a specular reflector, enabling strong double-bounce returns from tree trunks and branches (Hess *et al.* 1995). Polygon B, occupying a topographically lower position, is an open woodland which remained flooded at low water and had macrophyte cover and some open water at high-water stage. High-water backscattering decreased for polygon B because some of the trees and shrubs that were contributing to double-bounce scattering became submerged at high water. The cover at polygon C is seen on the high-water video frames to have roughly equal amounts of open water and aquatic macrophyte, with scattered trees;

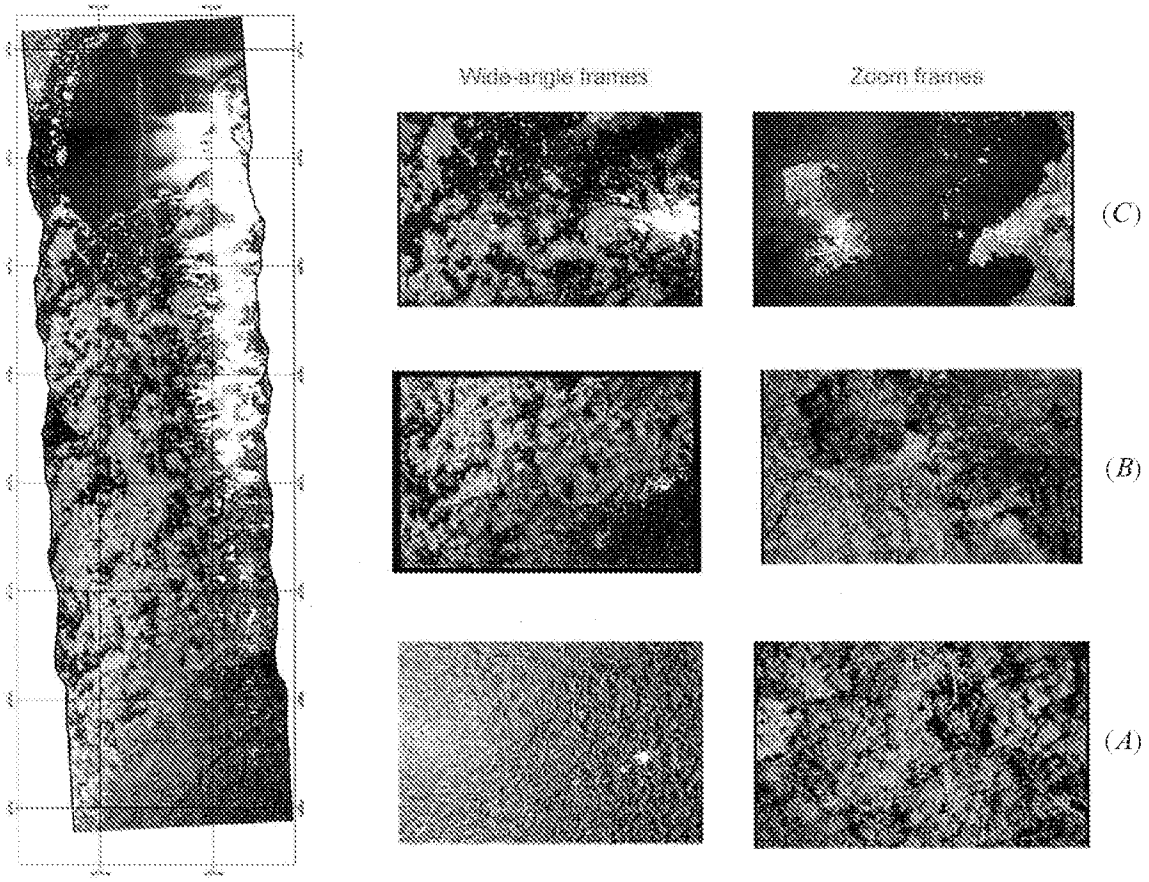


Figure 11. Mosaicked VOAM99 strip (30 s), and individual wide-angle and zoom frames at locations A, B, and C. Amazon floodplain near Lago Cabaliana, 11 June 1999. Transition from várzea forest on levee (*A*), through woodland savanna with floating macrophyte (*B*), to sparse woodland savanna with floating macrophyte (*C*), to open lake water.

at low water, the cover would be similar to that of polygon B at high water (and the backscattering coefficients confirm this). Backscattering decreases for polygon C at high water, as some of the emergent woody scattering elements are submerged and the per cent open water cover is higher. That unflooded forest (A, low water) as well as flooded open woodland and aquatic macrophyte (B, low water; C, high water) have a median σ° of -7.3 dB shows the importance of using multi-temporal data for wetland mapping with L-HH SAR data and the usefulness of videography for correctly interpreting the JERS-1 data to train classifiers.

VOAM95 and VOAM96 (non-digital) videography is being used to locate training sites for classes not well represented on the VOAM97 and VOAM99 surveys. Because water levels were 1–2 m higher during the 1997 and 1999 surveys than during the 1995 and 1996 JERS-1 imaging, the inundation states of areas such as lake beds and sand bars exposed at low water, or high levee forest that remained dry at high water, are best determined from the contemporaneous 1995 and 1996 videography.

6. Further applications of the VOAM data

6.1. *Location of ground control points in remote regions*

Effective use of geocoded video data as a validation dataset assumes the accurate geolocation of the satellite imagery with which the video data is being compared. Errors in geolocation of satellite imagery can significantly affect validation results for land cover mapping (Husak *et al.* 1999). Geocoding of the GRFM Amazon mosaics was accomplished by incorporating ground control points from 1:100 000 and 1:50 000 scale maps into the mosaicking process (Siqueira *et al.* 2000). While such maps exist for much of the Amazon basin, there is a coverage gap for a large portion of the western and north-western part of the basin, including much of the Manaus tile. Since the scale of the maps currently available for this region is too coarse for use in geocoding the mosaics, the map-based ground control points for the mosaic geocoding were supplemented with ground control points located on video mosaics created from VOAM99 survey data. An example of a ground control point used for the Manaus tile is shown in figure 12. Although relatively few features are both small enough to locate to within one to two pixels on the 100 m JERS-1 mosaics, and large and distinct enough to appear clearly on the mosaic, incorporation of a few such points can significantly increase the accuracy of geocoding for areas lacking accurate maps. With high-resolution optical imagery such as TM, matching of control points between satellite video images is easier, and the geocoded videography offers a practical means of building a database of ground control points in remote regions.

6.2. *High-resolution mapping of long-term study sites*

The high dynamic range and continuous automatic aperture and gain adjustments of digital video cameras allow them to adapt to rapidly changing light conditions and capture a reasonable representation of the ground even under a piebald sky. Good quality two-dimensional mosaics can thus be constructed from multiple flight lines acquired under widely varying light conditions. This capability was demonstrated during VOAM99 for the site of the Biological Dynamics of Forest Fragments Project (BDFFP), a collaborative research effort between Brazil's National Institute for Amazonian Research (INPA) and the Museum of Natural History of the Smithsonian Institution. The BDFFP site is also currently the focus of an LBA

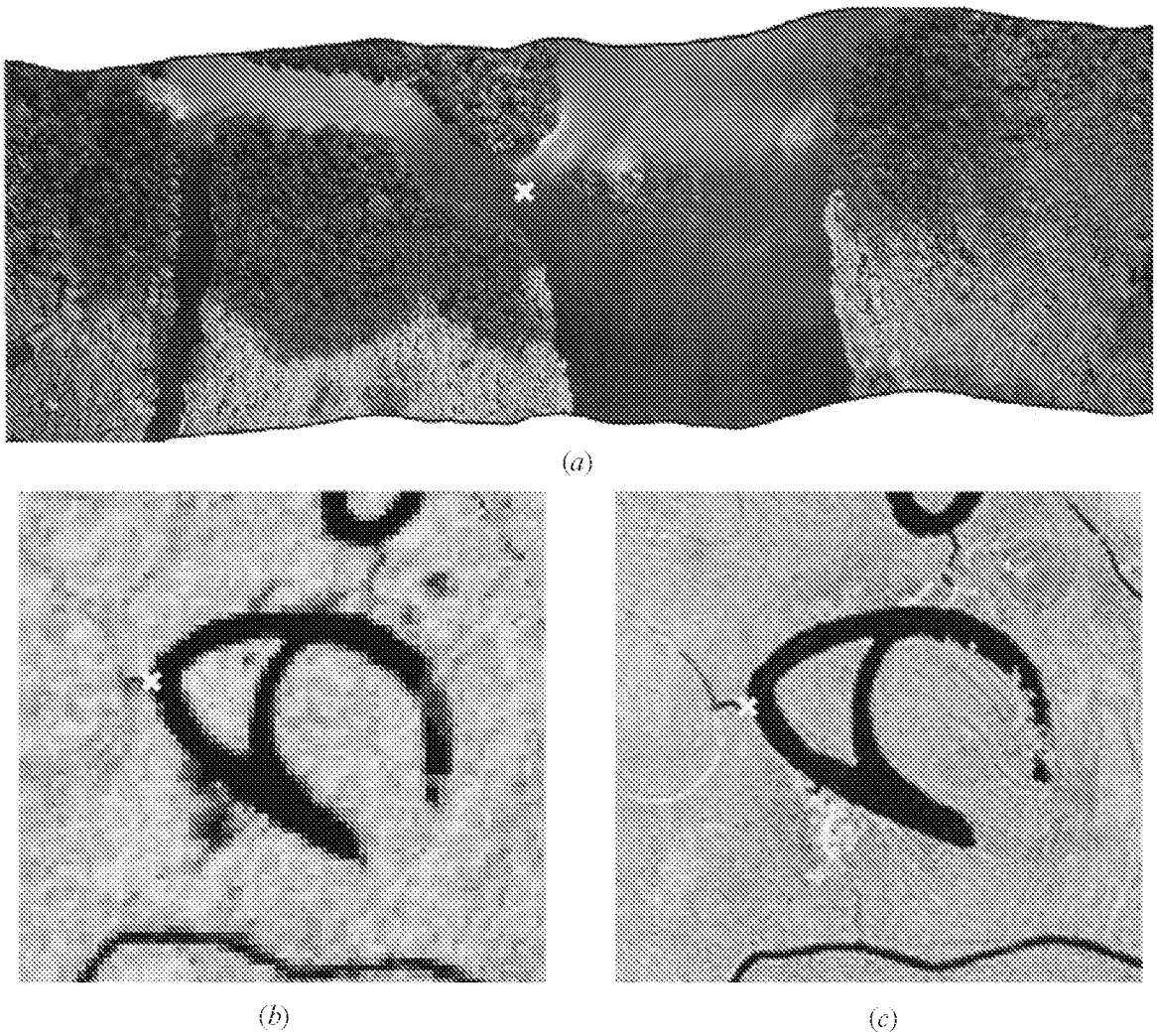


Figure 12. Control point at confluence of small channel and oxbow lake, located on (*a*) mosaicked video strip, (*b*) JERS-1 mosaic, and (*c*) Landsat TM scene. Japurá-Solimões floodplain (65.7974W, 2.2204S), 19 June 1999. Maps are not available for this region at scales larger than 1:250 000.

forest biomass investigation. Nearly complete coverage of the hilly site north of Manaus was acquired under broken cloud cover at 3000 m above ground. Over a period of 2.5 h, seven flight-lines were flown in a pattern that provided 50% side-lap for the wide-angle camera coverage of the 40×10 km area, and three of those lines were flown again at low altitude to obtain laser coverage. The higher-altitude video swaths were mosaicked using the methodology previously described. The georeferenced mosaic strips for each flight line were then rectified to each other and mosaicked together using the commercial GIS software TNTMIPS, producing a digital orthoimage with a resolution of 1.8 m and horizontal geolocation accuracy of approximately 6 m (Hayward and Slaymaker 2001). A portion of the orthoimage, which is being used as a research site map, is shown in figure 13.

6.3. *Forest biomass estimation*

The coverage of the BDFFP site is also being used in an ongoing project at the University of Massachusetts to develop a methodology for estimating standing biomass in tropical forests from aerial videography. A modification of the UMass

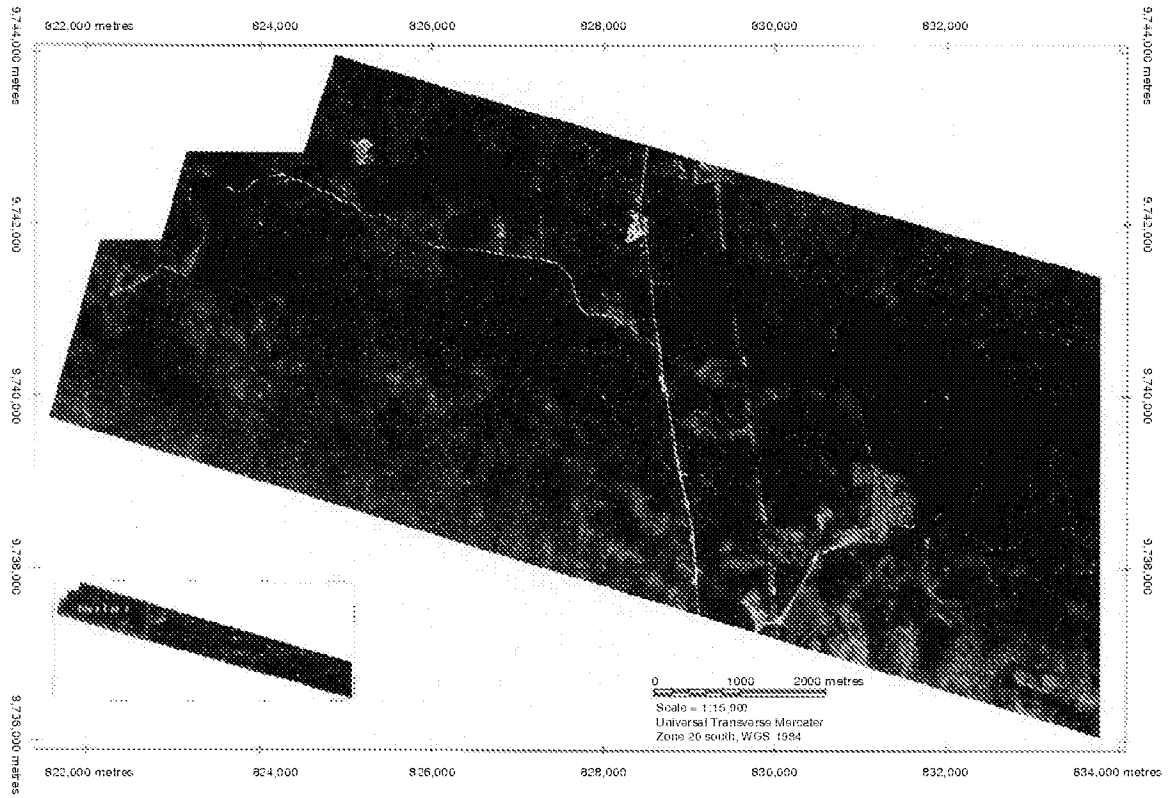


Figure 13. North-west portion of high-resolution digital orthoimage of BDFFP site, prepared from seven VOAM99 flight lines. Pixel dimensions 1.8 m.

mosaicking program is used to produce video strips in stereo pairs. Whereas the standard georeferencing program extracts a few lines from the centre of each video frame to construct its continuous swath, the stereo modification extracts two sets of lines from the leading and trailing portions of each frame, constructing separate mosaic strips that are a consistent epipolar pair. These strips can then be registered to the ground using the georeferenced centre strip and elevation data from the laser, viewed using software such as ERDAS Stereo Analyst, and used to make measurements in 3D space (Slaymaker *et al.* 1999).

The BDFFP study is based on 66 permanent, 1 ha study plots located in continuous forest and in forest fragments ranging from one to 100 ha in area. Tree crown diameters, tree heights, and diameter at breast height (dbh) of trees within each site have been measured, and their biomass has been calculated from allometric equations that use dbh as the significant measurement (Laurance *et al.* 1998). Several of these sites were traversed by low-altitude VOAM99 flight lines, with zoom coverage at 7 cm resolution. From these data, 3D model strips were constructed and used to estimate the above-ground biomass of the sites by calculating the average dbh per hectare (Hayward and Slaymaker 2001). Linear 1 ha plots were delineated across these sites on the zoom mosaics and best-fitting ellipses drawn in 3D space around every tree crown to calculate its average diameter. These crown measurements were then used to estimate the dbh values of those trees using a regression equation developed from BDFFP field data which correlated tree crown area to dbh. The average of the calculated dbh values for each site was compared to the actual average dbh for that site as determined from the BDFFP ground measurements, and good agreement was found between the average video-calculated dbh values and the

ground-measured values. Given the ease and speed with which aerial measurements can be made from the high-resolution 3D video model as compared to ground surveys, this approach may greatly increase the efficiency of monitoring programs that calculate regional biomass values to estimate the amount of sequestered carbon (Slaymaker *et al.* 1999).

6.4. *Rapid assessment of tropical forest fires*

The frequent repeat cycle of coarse-resolution sensors such as the DMSP (Defense Meterological Satellite Program) Operational Linescan System (OLS) makes them effective at monitoring large fires (Fuller and Fulk 2000). Fine-resolution monitoring is problematic in the tropics, however, because the longer repeat cycles of higher-resolution sensors, combined with the presence of smoke or clouds, can result in few satellite images suitable for rapidly estimating fire extent and damage. Geocoded videography was shown to be useful in an INPE multi-sensor assessment of damage to forests following the severe 1997–1998 fire season in the state of Roraima, Brazil (Shimabukuro *et al.* 1998). In that study, an equivalent to the VOAM97 system was flown on the INPE Bandeirante after the onset of the rainy season brought an end to widespread fires. Brazilian governmental agencies urgently needed to assess fire damage to forests, but cloud cover prevented the use of optical satellite imagery. Flying below the cloud ceiling of 1000–1200 m, the INPE plane sampled areas that had been identified by Brazilian Army fire crews and on DMSP-OLS imagery as having a high frequency of fires.

Interpreters were easily able to distinguish burned from unburned areas on mosaicked video strips, though automated interpretation was not possible owing to highly variable scene illumination caused by intermittent cloud cover. Two degrees of fire damage to forests could be determined from the videography: surface fires, which affected mainly the forest understorey; and canopy fires, which burned most of the overstorey as well as undergrowth and would require an extended period for forest recovery, possibly leading to significant changes in forest composition. Patterns of fire spread from agricultural clearings to surrounding forest could also be studied on the video mosaics. In the INPE study, the videography provided independent confirmation of estimates of burned forest area derived from Landsat imagery obtained several months after the end of the fire season.

7. Conclusion

Development of land cover validation datasets often involves a series of compromises, necessitated by limited financial resources as well as by physical constraints such as cloud cover, remote location, and dynamic ground conditions. Aerial videography helps overcome some of these limitations by providing a way to obtain high-resolution digital imagery from remote areas under cloud cover, at reasonable cost. When many hours of footage are collected, the ability to automatically geolocate and mosaic the video data becomes critical for effective analysis; the hardware and software described here for the VOAM97 and VOAM99 surveys were shown to meet this requirement.

The surveys documented changeable features such as inundation state under conditions the same as or similar to those prevailing during the GRFM Amazon mosaic acquisitions. Coverage was obtained for areas representative of the full range of wetland types found in the Brazilian Amazon, as well as for primary forest, forest clearings, and savannas, locatable to within approximately 10 m (VOAM99) or 100 m

(VOAM97) of the true location. Using the VOAM99 dataset, a methodology was demonstrated for acquiring a randomly selected set of validation samples for a frequently cloud-covered area extending 6° by 4° in longitude and latitude. Further applications demonstrated include obtaining control points for image geolocation, creating high-resolution digital orthoimages of intensive study sites, estimating forest biomass, and assessing fire damage to forests. To our knowledge, the VOAM surveys are the first to obtain a high-resolution, geocoded validation dataset for a subcontinental-scale tropical region. The resulting validation dataset is now being used by INPE and LBA investigators. We expect that it will not only serve as a validation source for wetlands mapping with SAR and optical data but will provide a benchmark for long-term monitoring of vegetation in the Brazilian Amazon.

Disclaimer

Trade names are given for the benefit of the reader and do not imply endorsement of particular products.

Acknowledgments

Invaluable technical support in assembling the video systems was provided by Bill Anderson, Dave Linden, Bob Pargee, Brent Burgess, and Tony Tonkin. Omnistar Inc. kindly provided a complimentary GPS satellite subscription. Bruce Forsberg and Ake Rosenqvist were instrumental in collecting the 1995 and 1996 videographic datasets. Special thanks to intrepid Bandeirante pilot Pedro Santos. Thank you to Bruce Nelson, Scott Anderson, Ivan Bergier, Rodrigo Lorene, Maycira Costa, and Reinaldo Barbosa for assistance during aerial and field surveys. For making the GRFM Amazon mosaics and VOAM surveys a reality, we thank Ake Rosenqvist, Masanobu Shimada, Diane Wickland, Tony Freeman, Bruce Chapman, Paul Siqueira, Greg McGarragh, Ryan Ashker, Howard Schultz, and Ed Riseman. The comments of three anonymous reviewers were very helpful. Support for this research was provided by NASA grants to GRFM-JAMMS studies at Jet Propulsion Laboratory (Tony Freeman) and UCSB (John Melack), and to LBA-Ecology investigation LC-07 (John Melack and Evlyn Novo). Acquisition of the JERS-1 mosaic imagery was made possible by NASDA's Global Rainforest Mapping Project.

References

Barbosa, C., Hess, L., Melack, J., and Novo, E., 2000, Mapping Amazon Basin wetlands through region-growing segmentation and segmented-based classification of JERS-1 data. *IX Latin American Symposium on Remote Sensing (SELPER), 6–10 November 2000*, in press.

Camara, G., Souza, R. C. M., Freitas, U. M., and Garrido, J. C. P., 1996, SPRING: Integrating remote sensing and GIS by object-oriented data modeling. *Computer Graphics*, **20**, 395–403.

Chapman, B., Siqueira, P., and Freeman, A., 2002, The JERS-1 Amazon Multi-Season Mapping Study (JAMMS): observation strategies and data characteristics. *International Journal of Remote Sensing*, 000–000.

Cihlar, J., 2000, Land cover mapping of large areas from satellites: status and research priorities. *International Journal of Remote Sensing*, **21**, 1093–1114.

Congalton, R. G., and Green, K., 1999, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices* (Boca Raton, Florida: Lewis Publishers).

Defries, R. S., and Belward, A. S., 2000, Global and regional land cover characterization from satellite data: an introduction to the Special Issue. *International Journal of Remote Sensing*, **21**, 1083–1092.

Eleveld, M. A., Blok, S. T., and Bakx, J. P. G., 2000, Deriving relief of a coastal landscape with aerial video data. *International Journal of Remote Sensing*, **21**, 189–195.

Everitt, J. H., Escobar, D. E., Summy, K. R., and Davis, M. R., 1994, Using airborne video, global positioning system, and geographical information system technologies for detecting and mapping citrus blackfly infestations. *Southwestern Entomologist*, **19**, 129–138.

Fuller, D. O., and Fulk, M., 2000, Comparison of NOAA-AVHRR and DMSP-OLS for operational fire monitoring in Kalimantan, Indonesia. *International Journal of Remote Sensing*, **21**, 181–187.

Graham, L. A., 1993, Airborne video for near-real-time vegetation mapping. *Journal of Forestry*, **8**, 28–32.

Hayward, C. D., and Slaymaker, D., 2001, Estimating the significant above ground biomass of Amazonian rain forest using low altitude aerial videography. *18th Biennial Workshop on Color Photography and Videography in Resource Assessment, 16–18 May 2001.*

Hess, L. L., 1999, Monitoring flooding and vegetation on seasonally inundated floodplains with multi-frequency polarimetric synthetic aperture radar. University of California Santa Barbara Santa Barbara, California.

Hess, L. L., Melack, J. M., Filoso, S., and Wang, Y., 1995, Delineation of inundated area and vegetation along the Amazon floodplain with the SIR-C synthetic aperture radar. *IEEE Transactions on Geoscience and Remote Sensing*, **33**, 896–904.

Husak, G. J., Hadley, B. C., and McGwire, K. C., 1999, Landsat thematic mapper registration accuracy and its effects on the IGBP validation. *Photogrammetric Engineering and Remote Sensing*, **65**, 1033–1039.

Irion, G., Junk, W. J., and Mello, J. A. S. N. D., 1997, The large central Amazonian river floodplains near Manaus: geological, climatological, hydrological, and geomorphological aspects. In *The Central Amazon Floodplain: Ecology of a Pulsing System*, edited by W. J. Junk (Berlin: Springer), pp. 23–46.

Jennings, C. A., Vohs, P. A., and Dewey, M. R., 1992, Classification of a wetland area along the upper Mississippi River with aerial videography. *Wetlands*, **12**, 163–170.

Jennings, R., 1999a, Consumer and professional digital video recording and data formats, http://www.oakmusic.com/parkplace/Video/DVPapers/dv_format.html.

Jennings, R., 1999b, Fire on the Wire: the IEEE 1394 High Performance Serial Bus, http://www.oakmusic.com/parkplace/Video/DVPapers/FireWire.html.

Junk, W. J., and Piedade, M. T. F., 1997, Plant life in the floodplain with special reference to herbaceous plants. In *The Central Amazon Floodplain: Ecology of a Pulsing System*, edited by W. J. Junk (Berlin: Springer), pp. 147–185.

Justice, C., Belward, A., Morisette, J., Lewis, P., Privette, J., and Baret, F., 2000, Developments in the 'validation' of satellite sensor products for the study of the land surface. *International Journal of Remote Sensing*, **21**, 3383–3390.

King, D. J., 1995, Airborne multispectral digital camera and video sensors: a critical review of system designs and applications. *Canadian Journal of Remote Sensing*, **21**, 245–273.

Laurance, W. F., Laurance, S. G., and Delamonica, P., 1998, Tropical forest fragmentation and greenhouse gas emissions. *Forest Ecology and Management*, **110**, 173–180.

Marsh, S. E., Walsh, J. L., and Sobrevila, C., 1994, Evaluation of airborne video data for land-cover classification accuracy assessment in an isolated Brazilian forest. *Remote Sensing of Environment*, **48**, 61–69.

Mausel, P. W., Everitt, J. H., Escobar, D. E., and King, D. J., 1992, Airborne videography: current status and future perspectives. *Photogrammetric Engineering and Remote Sensing*, **58**, 1189–1195.

Meisner, D. E., 1986, Fundamentals of airborne video remote sensing. *Remote Sensing of Environment*, **17**, 63–79.

Melack, J. M., Hess, L. L., and Sippel, S., 1994, Remote sensing of lakes and floodplains in the Amazon Basin. *Remote Sensing Reviews*, **10**, 127–142.

Muchoney, D., Strahler, A., Hodges, J., and LoCastro, J., 1999, The IGBP DISCover confidence sites and the System for Terrestrial Ecosystem Parametrization: tools for validating global land-cover data. *Photogrammetric Engineering and Remote Sensing*, **65**, 1061–1067.

Norton-Griffiths, M., 1988, Aerial point sampling for land use surveys. *Journal of Biogeography*, **15**, 149–156.

RICHEY, J. E., NOBRE, C., and DESER, C., 1989, Amazon River discharge and climate variability: 1903 to 1985. *Science*, **246**, 101–103.

RITTER, N., and RUTH, M., 1997, The GeoTiff data interchange standard for raster geographic images. *International Journal of Remote Sensing*, **18**, 1637–1647.

ROSENQVIST, A., SHIMADA, M., CHAPMAN, B., FREEMAN, A., DE GRANDI, G., SAATCHI, S., and RAUSTE, Y., 2000, The Global Rain Forest Mapping project—a review. *International Journal of Remote Sensing*, **21**, 1375–1387.

SAHAGIAN, D., and MELACK, J., 1998, Global wetland distribution and functional characterization: trace gases and the hydrologic cycle. Report from the joint GAIM/IGBP-DIS/IGAC/LUCC workshop, IGBP Report No. 46. IGBP Secretariat, Stockholm.

SCEPAN, J., 1999, Thematic validation of high-resolution global land-cover data sets. *Photogrammetric Engineering and Remote Sensing*, **65**, 1051–1060.

SCHAEFER, C. E. G. R., and DALRYMPLE, J. B., 1995, Landscape evolution in Roraima, north Amazonia: geomorphology, genesis and landscape evolution. *Zeitschrift fur Geomorphologie*, **39**, 1–28.

SCHULTZ, H., HANSON, A., HOLMES, C., RISEMAN, E., SLAYMAKER, D., and STOLLE, F., 1999a, Integrating small format aerial photography, videography, and a laser profiler for environmental monitoring. *ISPRS WG III/1 Workshop on Integrated Sensor Calibration and Orientation, 16–17 June.*

SCHULTZ, H., SLAYMAKER, D., HANSON, A., RISEMAN, E., HOLMES, C., POWELL, M., and DELANEY, M., 1999b, Cost-effective determination of biomass from aerial images. *International Workshop on Integrated Spatial Data, Portland, Maine, 14–15 June 1999,* published by Springer-Verlag, Berlin; pp 67–76.

SHIMABUKURO, Y. E., SANTOS, J. R. D., NOVO, E. M., KRUG, T., and HESS, L. L., 1998, Estimativa da Área de Cobertura Florestal Afetada pelo Incêndio em Roraima, Utilizando Dados de Multi-sensores. Instituto Nacional de Pesquisas Espaciais, São José dos Campos, Brazil.

SHIMABUKURO, Y. E., NOVO, E. M., and MERTES, L. A. K., 2002, Amazon River mainstem floodplain Landsat TM digital mosaic. *International Journal of Remote Sensing*, **23**, 57–69.

SIDLE, J. G., and ZIEWITZ, J. W., 1990, Use of aerial videography in wildlife habitat studies. *Wildlife Society Bulletin*, **18**, 56–62.

SIQUEIRA, P., HENSLEY, S., SHAFFER, S., HESS, L., MCGARRAGH, G., CHAPMAN, B., and FREEMAN, A., 2000, A continental-scale mosaic of the Amazon Basin using JERS-1 SAR. *IEEE Transactions on Geoscience and Remote Sensing*, **38**, 2638–2644.

SLAYMAKER, D. M., JONES, K. M. L., GRIFFIN, C. R., and FINN, J. T., 1996, Mapping deciduous forests in southern New England using aerial videography and hyperclustered multitemporal Landsat TM imager. In *Gap Analysis: A Landscape Approach to Biodiversity Planning*, edited by J. M. Scott, T. H. Tear and F. W. Davis (Bethesda, MD: American Society of Photogrammetry and Remote Sensing), pp. 87–101.

SLAYMAKER, D. M., SCHULTZ, H., HANSON, A., RISEMAN, E., HOLMES, C., POWELL, M., and DELANEY, M., 1999, Calculating forest biomass with small format aerial photography, videography and a profiling laser. *17th Biennial Workshop on Videography and Color Photography in Resource Assessment, April 1999* (Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing), pp. 241–260.

STEHMAN, S. V., and CZAPLEWSKI, R. L., 1998, Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sensing of Environment*, **64**, 331–344.

STRONG, L. L., and COWARDIN, L. M., 1995, Improving prairie pond counts with aerial video and global positioning systems. *Journal of Wildlife Management*, **59**, 708–719.

TORTORA, R., 1978, A note on sample size estimation for multinomial populations. *American Statistician*, **32**, 100–102.

ZHU, Z., RISEMAN, E. M., HANSON, A. R., and SCHULTZ, H., 1999, Automatic Geo-Correction of Video Mosaics for Environmental Monitoring.

IMAGE PROCESSING IN THE GRASS GIS

Hinthorne, J., Satnik, D. (1), Vali, A., and Kleeman, T. (2)

(1) GIS Lab, Central Washington Univ., Ellensburg, WA 98926.
(2) RIA Inc., 317 RR 620 South, Suite 302, Austin, TX 78734.

ABSTRACT

The UNIX-based public domain raster GIS named GRASS has some excellent user friendly programs for the import and export of data, and several kinds of image processing for band files and/or maps including: spatial filtering; arithmetic combination (band ratios, etc.); Fast Fourier transformation, band space transformations (principle components, canonical components, etc.); classification (supervised and unsupervised); and rectification or georeferencing. All raster processing programs can be applied to all raster images and maps in GRASS data bases; all use the low level GRASS library to insure portability and future compatibility; and all are executed by commands using standard GRASS and UNIX syntax.

image processing, GRASS, geographic information systems, GIS

GENERAL DESCRIPTION OF GRASS

Since 1983 the UNIX-based GRASS (Geographic Resources Analysis Support System) geographic information system, developed and maintained by the Construction Engineering Research Laboratory (CERL) of the U. S. Army, has evolved into a very powerful, public domain, raster based GIS with a full range of single-layer and multi-layer map analysis tools. It also contains a number of import/export and map drawing functions which accommodate vector and point data types, as well as raster type. Recently, several programs have been added to integrate the GRASS user's environment with a relational data base management system for point and vector data types.

IMAGE PROCESSING

Some image processing functions have been available to GRASS users since the development of version 1 in 1984, however these were contained in an external "GRASS-IMAGERY" software module for data from satellite and photographic sources only; data layers developed in the main "GRASS-GRID" module could not be processed with the image enhancement and classification routines. Today, in versions 3 and 4, all raster analysis functions, including those traditionally viewed as "image processing," are contained in a single comprehensive set of software tools which use the same low-level library functions for file access and graphics display and are executed by commands with consistent user interfaces.

GRASS image processing functions can be grouped into seven categories: import / export, spatial domain transformations, frequency domain transformations, band/color space transformations, classification and rectification. All functions can be used on any raster image or map in any logical order; e.g., classification can be used either before or after rectification. In addition, any non-"image processing" analysis function can be used on image bands when appropriate. Individual cell values may be up to 4 byte integers in most current workstation imple- mentations of GRASS, and disk storage is automatically reduced by a run length encoding algorithm.

Program names are marked by italics and programs marked with the † symbol have been recently added to GRASS by us.

IMPORT / EXPORT

GRASS has routines to automatically read Landsat Multi-Spectral Scanner (MSS) and Thematic Mapper (TM) data. TM distribution tapes may be in either band sequential or band interleaved formats. SPOT imagery, NHAP (National High Altitude Photography) and other "raw raster" data formats are all imported from

CH2971-0/91/0000-2227$01.00 © 1991 IEEE

tape or CD-ROM by a routine which queries the user for necessary information such as tape block size and the number of bytes per cell. UNIX utility programs are used to accomplish functions such as byte swapping and removal of unnecessary end of line characters. Raster images can also be input from text files (ASCII) where each row of an image is a string of space separated numbers with new-line characters marking the end of each row. Vector data in a number of standard formats can be imported into GRASS and converted raster maps.

Because the data compression routine in GRASS can be processor dependent, for export the *decompress* function is normally run to produce a binary file of the size rows*cols*bpc bytes (where bpc is the number of bytes per cell required by the largest cell value in the image). The resulting file can then be exported via tape, modem, etc., using UNIX tools to transfer the data to another GRASS site or other image processing/GIS system. Supporting information such as color tables, descriptions of cell values, histograms and geographic location data are stored in GRASS as text files and may be transferred directly to other sites or printed. GRASS also has routines to convert raster images to vector maps and these can be exported in a variety of formats, e.g., DLG-3, ARC/INFO, AUTOCAD-DFX.

SPATIAL DOMAIN TRANSFORMATIONS

Normal image filtering for smoothing, feature enhancement, etc., can be accomplished by several functions in GRASS. *Gmfilter* performs the standard function of convolving an odd sized square matrix of user-defined factors to an image or cell map, summing the products and dividing by a user-defined value to calculate a new value for each cell in an output image. A choice between sequential or parallel processing of cells is given. The *neighbors* program also works on the basis of a matrix-like neighborhood around each cell, but offers a choice of several filter algorithms, some of which cannot be accomplished with *Gmfilter*, e.g., mode, interspersion, maximum, minimum, or diversity.

General purpose arithmetic computations can be done on all GRASS raster maps and images with *Gmapcalc*. This function provides a combination of arithmetic (addition, subtraction, multiplication, division and modulus) operators, logical (and, or, not, and if-then-else) operators and transcendental functions (trigonometric, exponential and logarithmic). Operands for the calculations can be constants (integer or floating point), cell values from the image or map, or logical constants (true is 1 and false is 0). Up to 15 maps and/or images can be used in a single *Gmapcalc* expression to produce a new image or map. Each cell, row by row, is calculated in turn. Two examples will indicate the power of *Gmapcalc*.

To compute a new image which is the average of three band file images:

Gmapcalc aver.123=(band1+band2+band3)/3

Or, to produce an elevation model in feet from one in meters:

Gmapcalc elev.ft=elev.meters * 39.4 / 12.0

Recent additions to *Gmapcalc* provide for the optional specification of a cell on each input map relative to the current one being computed (offset by a number of rows and/or columns) to be used as an operand. This makes it possible to calculate slope and aspect maps from digital elevation models or perform the image filtering like *Gmfilter* with greater flexibility (exponential or logarithmic scaling for example). Very complex modeling calculations can be implemented using *Gmapcalc*.

Gmapcalc is commonly used to calculate band ratios to enhance some features of interest in an image. And can be used to perform many other types of image enhancement.

FREQUENCY DOMAIN TRANSFORMATIONS

Many image defects such as bad or missing scan lines or periodic noise are difficult to remove with filters in the spatial domain, but are dealt with more easily in the frequency domain.

A Fast Fourier Transform (FFT) program, *Gfft*† and its inverse *Gifft*†, have recently been added to GRASS by us. The precision required to perform the inverse FFT is maintained by saving double precision floating point copies of the real and imaginary components of the FFT as well as creating scaled integer real and imaginary component images for viewing, masking, etc., by the user. Rectangular images can be processed by *Gfft* and *Gifft*; they are zero-filled in each direction to create a rectangle whose sides are an even power of two. The transformation is computationally intensive, but our implementation uses an algorithm of order N $\log_2$ N originally devised by N. Brenner of Lincoln Laboratories.

Using the FFT routines, a zero-crossing edge detection filter has been implemented, *Gzc*†, based on the Laplacian of the gaussian function. The user can control the width and threshold of the gaussian filter as well as the number of directional (azimuthal) categories in the product image.

High and low pass frequency filters can be performed on an image band by using *Gfft* to transform the image into the frequency domain, then using the standard GIS masking function to mask out the high or low frequency components and finally using *Gifft* to transform the masked frequency domain image back into the spatial domain.

BAND/COLOR SPACE TRANSFORMATIONS

The following band and color space transformations have been recently added to GRASS by us. The first two band space transforms, *Gpca*† and *Gcca*†, depend on statistical information taken from the image to determine the parameters of the transformation. The third transformation, *Grgb.to.his*†, treats the image bands as color components and performs a color space transformation.

The program *Gpca* has been developed to compute N (where $2 \leq N \leq 8$) principle component image bands from N input image bands. Principal components analysis uses a linear combination of the bands to transform a set of correlated bands into a set of uncorrelated bands. This process condenses the information content of the multiple input image bands in the first few output image bands, typically leaving noise-like data in the last few output bands.

The *Gcca* program performs a canonical components transformation on N (where $2 \leq N \leq 8$) input image bands and a set of training regions to compute N output image bands. Canonical components analysis uses a linear combination of the input image bands that maximize the separability of the training regions to generate the set of output image bands.

The *Grgb.to.his* program (and its inverse *Ghis.to.rgb*†) perform color space transformations on the image bands. Each of these programs take three input image bands and generate three output image bands. *Grgb.to.his* treats the three input bands as the red, green and blue components of the image and produces the three output bands as hue, intensity and saturation.

IMAGE CLASSIFICATION

The GRASS programs *i.cluster* and *i.class*† provide for the generation of sets of class (category) signatures from up to 8 input bands. In *i.class* the user graphically selects the training area (supervised classification) to define each class and is then shown all the regions of the image that match the signature using a parallelapiped classification method. *i.cluster* generates a set of signatures by an "unsupervised" method; the user can specify the maximum number of classes and the percent convergence in the iterative assignment of cells to classes. Seed signatures for *i.cluster* can be taken from *i.class*, previous *i.cluster* results or entered by hand.

A single "classifier" program, *i.maxlik*, processes the signatures and the image band files to create a final classified image using a maximum likelihood method. Other classifier algorithms could be added to GRASS by users with C programming experience. (The complete source code of GRASS is distributed to all users.)

RECTIFICATION

The conversion of raw images into geographic coordinate systems, known as rectification or georeferencing, combines translation, rotation, scaling and differential stretching. Three GRASS programs are used to accomplish this. The first is *i.target* in which the user identifies the data base into which the rectification will be done; this "target" data base is typically one with maps in the UTM coordinate system. Second, the *i.points* program is used to associate "true" ground coordinates with a number of specific cells or points on the image(s) to rectify; typically 15 to 30 points are desired for superior results. Then the *i.rectify* program is run which rectifies designated image bands and/or processed image band products into the target data base.

CONCLUSIONS

Image processing using the GRASS GIS offers many possibilities for enhancement and transformation of image bands. The versatility of programs like *Gmapcalc* and *neighbors* coupled with the special purpose commands like *Gpca* and *Gcca* provide a rich set of tools for the user. In addition, because of the UNIX operating environment underneath GRASS the user and/or programmer can readily build shell scripts (batch files) that allow a series of the image processing commands to be executed as a single command. This leads to an extendible and user customizable environment which is quite powerful.

CHAPTER XVIII

Mapping from Oblique Photographs

Author-Editor: RALPH K. IRBOP
Contributing Author: RUSSELL C. DOOLITTLE

18.1 Introduction

AN OBLIQUE photograph is defined as a photograph taken with the camera axis directed intentionally between the horizontal and the vertical. A further expedient differentiation has been made between high-obliques and low-obliques by defining high-obliques as oblique photographs in which the apparent horizon is shown, and low-obliques as those in which the apparent horizon is not shown because of the low tilt. This chapter will deal primarily with mapping from high-obliques and particularly with high-obliques as exemplified by trimetrogon photographs.

For many years, oblique photography was used for small-scale mapping because it offered the following advantages:

A. A single strip covers a wide strip of terrain.
B. The requirement for highly precise flying of photography is relaxed.
C. Variations in altitudes, directions, and spacing of flights do not form serious compilation hazards.
D. The need for geographic control is greatly reduced.
E. The requirements for personnel, aircraft, cameras, and film per unit ground area are greatly reduced.

In comparison with verticals obtained from the same altitude, oblique photographs possess the following disadvantages:

A. It is comparatively difficult to select, identify, and transfer control points on oblique photographs.
B. It is comparatively difficult to identify and delineate significant cartographic features on oblique photographs. The oblique view results in the "hiding" of features behind relief.
C. The procedures for utilizing obliques involve the use of several operations not required for vertical photographs. The present trend is definitely toward much more extensive use of low-obliques, particularly the convergent type, and verticals. The primary factor influencing this trend has been the advent of the ability to obtain photographs from very high altitudes. This has greatly reduced the significance of the advantages of oblique photography.

18.2 Trimetrogon Photography

Trimetrogon aerial photography has been used extensively by the U. S. Air Force for the rapid compilation of small-scale aeronautical charts. However, no significant trimetrogon compilation has been accomplished by the U. S. Air Force since 1956 or by the Canadian Department of National Defence since the completion of the Canadian 8-mile series in 1957. In general, it can be said that, while trimetrogon compilation is no longer being used extensively by any major mapping agency, the use of trimetrogon photography typifies the problems encountered when using any type of oblique photography for small-scale mapping purposes.

18.2.1 GENERAL

In the United States the tricamera system of map compilation is designated as the trimetrogon method. The word Metrogon is the commercial name of the type of lens generally used, but any wide-angle lens of suitable quality may be used. This type of photograph is obtained by an assembly of three cameras, one pointing vertically downward in the conventional manner and two mounted obliquely, pointing in directions which will be as nearly perpendicular to the line of flight as possible. The two obliquely mounted cameras are oriented so that they will photograph

OFEK07693

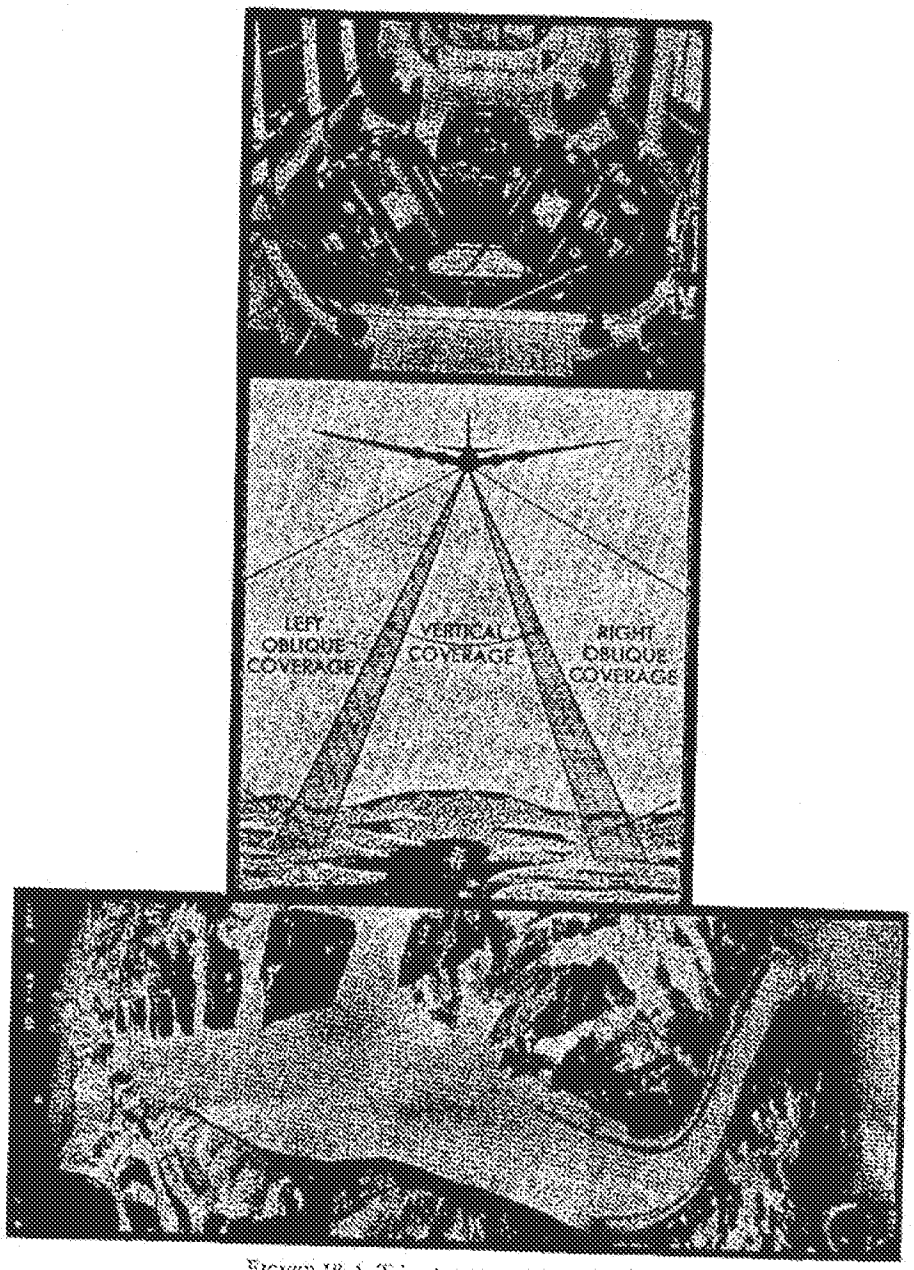


FIGURE 18-1. Trimetrogon camera assembly.

both the horizon and a small part of the area covered by the vertical camera. All three cameras are exposed simultaneously and thus photograph a strip of ground extending from horizon to horizon in a direction perpendicular to the flight line (see fig. 18-1).

18.2.1.1 CAMERAS

Either the K-17 or K-38 types of cameras are used in the trimetrogon assembly. In either case, they employ a 6-inch cone equipped with the Metrogon wide-angle lens and an A-5 vacuum-back magazine exposing a 9- by 9-inch negative. These

OFEK07694

cameras have an angular coverage of approximately 74 degrees as measured across the axis of the focal plane. Thus three of them provide more than enough coverage for the required 180 degrees from horizon to horizon. This mounting provides for an overlap of approximately 14 degrees between the fields of coverage of the oblique and vertical cameras. With the aircraft in level flight, the oblique cameras cover approximately 7 degrees above the horizon.

It is desirable to have the cameras precisely calibrated so that lens distortion and the calibrated focal length are known. The cameras are equipped with filters to reduce the effects of haze and improve definition of the apparent horizon. A vignetting filter is also used to produce negatives of uniform density. The operation of the three cameras is controlled by a single intervalometer, so that all are actuated simultaneously.

18.2.1.2 Specifications

Maximum efficiency for compilation is attained when the flight lines are straight and parallel. It is important that all three cameras be in operation continuously and that the horizon always appear in the obliques. Adherence to the following specifications will produce optimum-quality photography.

Flight height.—The flights should be at 20,000 feet or more above the mean elevation of the terrain, insofar as the service ceiling of the aircraft will permit. The time required for compilation per unit area varies inversely with the flight height up to about 40,000 feet. Since trimetrogon photography is used only for charting at scales of 1:100,000 or smaller, flight heights below 20,000 feet are uneconomical and should be avoided except when it is necessary to fly below cloud cover. There should not be a difference in height above mean terrain of more than 5,000 feet between the highest and lowest flight lines in any one block of photographs. Care should be taken to maintain the specified altitude within each flight and in no case should deviation therefrom be more than 250 feet.

Overlap.—The end (or forward) lap should be between 55 and 65 percent. Because the trimetrogon photography gives horizon-to-horizon coverage, the side lap between flights is specified in terms of flight-line spacing, instead of in percentages, as is customary for vertical photography. Recommended flight-line spacings for preparation of maps at scales of 1:250,000 and 1:1,000,000 are given below. These should not be exceeded by more than 10 percent.

Flight height	*1:250,000*	*1:1,000,000*
20,000 feet	15 miles	20 miles
30,000 feet	20 miles	25 miles
40,000 feet	22 miles	28 miles

Where flight lines parallel the coast, the flight line adjacent to the coast will be flown inshore as far as possible while still including the beach on the vertical photographs. Insofar as possible, the principal points of no two consecutive verticals should be over water. In accomplishing coastline strips, it is desirable that all three cameras be operated, even though it might seem that the seaward-pointing camera is not needed, because knowledge of the presence or absence of islands, reefs, rocks, or shoals offshore is important and the horizon image is extremely useful for tilt determination.

Time of photography.—Photographs should be taken on days of clear weather, and preferably during the part of the day when the sun is 30 degrees or more above the horizon. In polar latitudes and in areas where there are fogs or clouds during the middle of the day for weeks on end, photographs may be taken at any time when the solar light is sufficient to get adequate detail in the shadows. Care should be taken to have minimum haze and cloud conditions extending to 25 miles on each side of the flight. It is preferable to have minimum cloud conditions extend to the apparent horizon, but the presence of clouds more than 25 miles from the flight line will not preclude the taking of photographs.

Flight data.—For filing and making use of trimetrogon photographs, certain information supplied by the camera crew should be shown on the negatives. Exposures at the ends of strips and at the ends of rolls are key exposures and should show the exposure number, camera identification (left, vertical, or right), mission (flight or roll number), project designation, date, focal length of lens (in millimeters), flight height, and geographic coordinates. Photographs between the key exposures should show the exposure number, camera identification, mission (flight or roll number), and project designation.

OFEK07695

18.2.2 PLANNING

Because of the large amount of material and the numerous operations involved in the usual trimetrogon compilation project, it is advisable to formulate a concise plan of procedure to serve as a guide to the various compilation units. This plan should take into account the type and scale of the map to be produced, the quality of the trimetrogon photography, and the availability and reliability of geographic control and source maps.

Indexes showing the location and coverage of all the photographs are made, usually as overlays to 1:1,000,000-scale maps. These indexes should show the location and number of every fourth or fifth vertical with indications of flight directions, the beginning and end of each flight, and notes as to missing prints, etc. Normally, the scale and coverage of the trimetrogon photographs are such that only alternate sets (a vertical and its right and left obliques constitute a set) of photographs are utilized in the compilation of the map.

Planning diagrams are then made showing the size, type, and scale of the map projection to be used for the templet assembly; the location of the geographic control, adjoining sheets, and source maps; and the location of the flights selected for compilation, with notations as to the photographs to be used and the points where azimuth transfers or ties are required at flight junctions. The scale should be indicated for the map projection to be used for the templet assembly, and, in turn, for the manuscript base. The scale should be carefully selected so as to be commensurate with the scale of the photographs, final map accuracy required, and the area available on which to lay out the base. A scale larger than the scale of the photographs should not be selected. When using photographs taken according to the specifications listed in paragraph 18.2.1.2 of this chapter, a manuscript scale of 1:80,000 to 1:100,000 will be the most satisfactory.

18.2.3 CONTROL POINTS

Photogrammetric control points are strategically located image points which appear on each of several overlapping photographs and which can be fixed in position on the manuscript by radial triangulation and used to position on the manuscript the planimetric detail appearing on the photographs. These points should be selected with the aid of a stereoscope to insure their accurate identification and transfer. Because of the large number of photogrammetric control points used in the trimetrogon method, it is convenient to subdivide them into the following classes which are listed in the order in which they are usually selected and marked.

Tie points are used to maintain common scale between two adjacent flight strips and to link them together. They should be placed so that they will be equally spaced along a line equidistant between the two flight lines. These points will usually fall near or slightly beyond the principal points of the obliques, depending on the flight spacing. If possible, a sufficient number of points should be selected so that there will be three well spaced points appearing on each oblique for which a templet will be made.

Pass points are used to tie together the adjacent photographs in a single flight strip. The tie points usually serve as the outer pass points, but in cases where it is impossible to pick tie points, a set of pass points should be picked in the same positions that tie points would usually occupy. The inner set of pass points should be picked in the area of overlap of two verticals and their corresponding obliques. Inasmuch as alternate sets are normally used, the overlap between adjacent verticals will usually be 10 to 20 percent or 1 to 2 inches for 9- by 9-inch prints.

Annex points are used to help in the relative orientation of the vertical and oblique prints. They are picked in the overlap between the vertical and its corresponding obliques about midway between the two inner pass points. Since only alternate sets of photos are used, each annex point will have only one radial.

Detail points are supplemental control points picked on the oblique photos at or near the important planimetric, cultural, and topographic features selected for compilation. The number of detail points selected will vary with the amount of detail required for the chart being compiled and the nature of the terrain. An average of 10 detail points per photograph will normally suffice.

Nadir points and azimuths are plotted on the vertical photographs by offsets from the principal points, the amount and direction of the offsets being determined by tilt

OFEK07696

analysis as described below. The radials connecting the nadir points of adjacent photographs are called "azimuths." It is usually impossible, because of insufficient overlap, to transfer a nadir point from a photo to the next alternate photo by direct image match. The azimuths, as a result, must be plotted as follows: (1) Transfer the nadir points of two adjacent verticals to the unused intermediate photo by image match; (2) connect the transferred points by a finely ruled line; (3) select several well spaced images which fall on this line on the intermediate photo and with the aid of a stereoscope prick the corresponding images on the photographs to be used; (4) draw the azimuths as straight lines through the nadir points and the pricked image points.

18.3 Tilt Analysis

The tilt of each photograph with respect to a horizontal plane must be found so that the photograph can be rectified to obtain the true horizontal angles between the nadir point radials.

The tilt analysis of trimetrogon photographs is greatly facilitated if the three cameras are rigidly mounted with a known geometrical relationship, because tilt values determined for one of the three photographs in a set can then be used to determine the tilt values of the other two photographs in the same set. Usually it is necessary to compute these values for only a few sets of photographs in any one strip; but if the relative positions and orientations of the cameras are not rigidly fixed or if the cameras are not exposed simultaneously, it may be necessary to check each set. In order to determine the exterior orientation of the photographs with respect to a horizontal plane, the accidental tilts caused by the pitch and roll of the aircraft must be analyzed. This task is not as formidable as it would seem, because it is usually necessary to make complete tilt computations for only a few sets, the values on intermediate sets being determined by transfer methods. In fact, it is frequently possible to adopt average tilt values for an entire flight. The orientation of the vertical photograph is expressed in terms of the components of tilt about a set of rectangular axes; the orientation of the oblique photograph is expressed in terms of the total tilt and the swing (direction of the principal line).

18.3.1 DEFINITIONS

Figure 18–2 shows a perspective view of the three positive prints from a trimetrogon assembly with the common exposure station. Figure 18–3 shows the vertical and right oblique prints in a tilted position.

Listed below are definitions of several terms which are used only in trimetrogon compilation:

Break angle.—The deflection angle (smaller dihedral angle) between the two vertical planes passing through the common exposure station and the principal points of the left and right obliques.

Depression angle.—The vertical angle θ between the axis of the oblique camera and the horizontal plane representing the true horizon. The depression angle is the complement of the angle of tilt.

Dip angle.—The angle in the principal plane between the horizontal plane representing the true horizon and the ray from the exposure station to the apparent horizon.

Interlocking angle.—The angle between the optical axes of the vertical and oblique cameras $(A+\beta)$ $(T+\Delta)$. The dihedral angle between the planes of the vertical and oblique photographs.

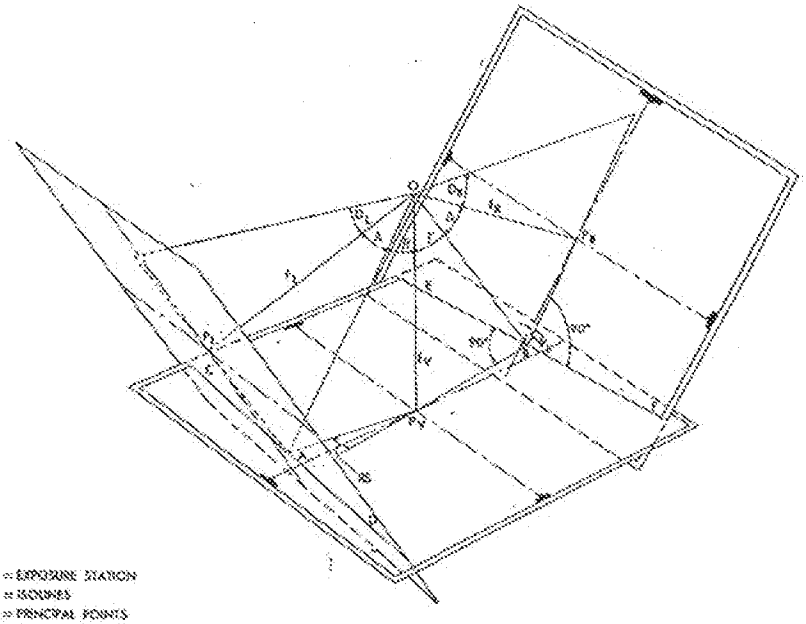
Isoline.—A line representing the intersection of the planes of the vertical and oblique photographs (CD, EF, in fig. 18–2). If both photographs have the same principal distance, the isoline will also be a line of common scale.

Locking angle.—The complement D of the interlocking angle. The depression angle of the oblique when the tilt of the vertical photo is zero.

Relative setting.—The dihedral angle RS between the two planes passing through the principal points of the opposite obliques, the principal point of the vertical photograph, and the common exposure station. This angle is measured on the vertical photograph as the angle between the two isolines, or as the deflection angle between the perpendiculars from the principal point of the vertical photograph to the two isolines.

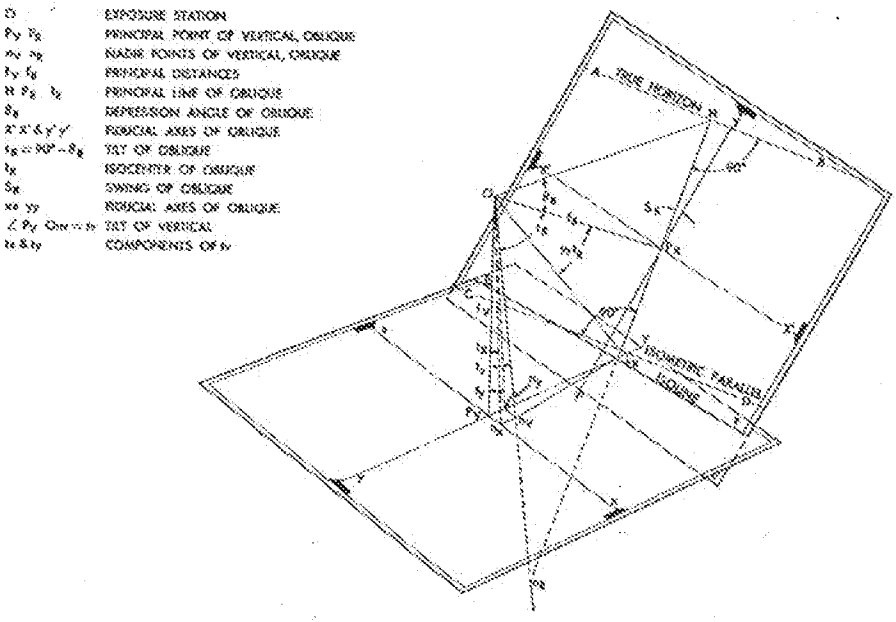
Relative swing.—The angle of rotation Ψ of the oblique camera about its own axis

OFEK07697



O = EXPOSURE STATION
CO AND CR = ISOMINES
P_L, P_V AND P_R = PRINCIPAL POINTS
D_V AND D_R = TOCKING ANGLES
$(A+B)$ AND $(A-B)$ = INTERLOCKING ANGLES
f_L, f_V AND f_R = PRINCIPAL DISTANCES
Φ_R = RELATIVE SWING
RS = RELATIVE SETTING

FIGURE 18–2. Geometry of the trimetrogon assembly.



O = EXPOSURE STATION
P_V P_R = PRINCIPAL POINT OF VERTICAL, OBLIQUE
n_V n_R = NADIR POINTS OF VERTICAL, OBLIQUE
f_V f_R = PRINCIPAL DISTANCES
N P_R t_R = PRINCIPAL LINE OF OBLIQUE
δ_R = DEPRESSION ANGLE OF OBLIQUE
x'x' & y'y' = FIDUCIAL AXES OF OBLIQUE
$t_R = 90° - \delta_R$ = TILT OF OBLIQUE
i_R = ISOCENTER OF OBLIQUE
S_R = SWING OF OBLIQUE
xx yy = FIDUCIAL AXES OF OBLIQUE
$\angle P_V O n_V = t_V$ = TILT OF VERTICAL
x_R & y_R = COMPONENTS OF t_R

FIGURE 18–3. Vertical and right oblique in a tilted position.

OFEK07698

with respect to the plane of the vertical photograph, measured on the oblique photograph by the angle between the isoline and a line joining the fore and aft fiducial markers.

Swing.—Generally, the rotation of a photograph in its own plane about its camera axis. On trimetrogon obliques, the angle between the principal line and the y axis, or the angle between the isometric parallel and the x axis.

18.3.2 RELATIVE CAMERA ORIENTATION

The elements of relative camera orientation are the interlocking angles, the relative swing, and the relative setting. These angles are all determined from the position of the isolines on the vertical photo and the corresponding obliques. The procedures for their determination is given below.

Isolines.—Because the focal lengths of the three cameras seldom vary by more than 1 percent, they may be regarded as equal and used to determine a substitute isoline parallel to the true isoline.

1. Using a compass describe arcs of the same radii (approximately 4.5 inches) from the principal points to both corners of the area of common overlap on both the vertical and one of the corresponding oblique photographs.
2. In each corner of the overlap area, locate the common image point which falls on the inscribed arcs on both the vertical and oblique photographs. These points correspond to E and F in figure 18-2.
3. Repeat steps 1 and 2 using arcs of 4-inch radius.
4. Etch straight lines joining the four common image points on both the vertical and oblique photos. All images intersected by the isoline on the oblique photo must likewise be intersected by the corresponding isoline on the vertical photo.
5. Repeat steps 1 through 4 using the vertical and the opposite oblique photograph.

Interlocking and locking angle.—The interlocking angle between the vertical and the corresponding oblique photographs is determined by the method of tangents as follows:

1. On the oblique photo, drop a perpendicular to the isoline from the principal point.
2. Near the intersection of the isoline and the perpendicular, select an image point common to the oblique and vertical photographs. The image should be within 0.05 inch of the perpendicular and as close to the isoline as possible.
3. Measure the distances to this image from the principal points of the vertical and oblique photographs and divide each distance by the appropriate principal distance.
4. From table of natural tangents, obtain angles corresponding to the two values obtained in 3 and add to get the interlocking angle.
5. Obtain the locking angle D by subtracting the value obtained from 90°.

Relative swing.—Proceed as follows:

1. On the oblique photograph, draw a line parallel to the x axis (line joining the fore and aft fiducial markers) intersecting the isoline near the aft edge of the photograph.
2. Measure the relative swing as the angle between the x axis parallel and the isoline.
3. Record relative swing as plus (+) if the forward end of the isoline is above the x axis parallel; record as minus (−), if below.

Relative setting.—Proceed as follows:

1. Measure the relative setting as the angle between the two isolines on the vertical photograph.
2. If the intersection of the two isolines (produced) is aft of the vertical photo, the relative setting is recorded as plus (+). If the intersection is forward, the relative setting is recorded as minus (−).

Recording data.—In each flight, the relative orientation should be determined and recorded for at least three sets of photographs, one each at the beginning, middle, and end of the strip. If the values for the three sets agree within 10′ of arc, the averages of the three determinations are used for the entire strip. If it is indicated that the camera mountings were not rigid, the relative orientation of every set must be computed.

18.3.3 EXTERIOR ORIENTATION

The elements of exterior orientation to be determined are the depression and swing angles of the oblique photographs, the x tilt and y tilt of the vertical photos, and the break angle between the principal planes of the two obliques in each set. Tilt determination is accomplished using the apparent-horizon technique. The steps are as follows:

A. Apparent horizon
   1. If the image of the apparent horizon is clear on the photograph so that the apparent horizon appears as a curve, select one point at each end of this curve and draw a line connecting the points selected. When the quality of photography is such that the apparent horizon cannot be clearly defined, more points may be necessary in order to best define the curve.

OFEK07699

TABLE 18-1.—RELATIONSHIP BETWEEN DIP AND ALTITUDE. DIP (MINUTES) $= 0.98\sqrt{H}$ (FEET)

Altitude	*Dip*	*Altitude*	*Dip*	*Altitude*	*Dip*	*Altitude*	*Dip*
1,000	0°31′	11,000	1°43′	21,000	2°22′	31,000	2°53′
2,000	0°44′	12,000	1°47′	22,000	2°25′	32,000	2°55′
3,000	0°54′	13,000	1°52′	23,000	2°29′	33,000	2°58′
4,000	1°02′	14,000	1°56′	24,000	2°32′	34,000	3°01′
5,000	1°09′	15,000	2°00′	25,000	2°35′	35,000	3°03′
6,000	1°16′	16,000	2°04′	26,000	2°38′	36,000	3°06′
7,000	1°22′	17,000	2°08′	27,000	2°41′	37,000	3°09′
8,000	1°28′	18,000	2°11′	28,000	2°44′	38,000	3°11′
9,000	1°33′	19,000	2°15′	29,000	2°47′	39,000	3°14′
10,000	1°38′	20,000	2°19′	30,000	2°50′	40,000	3°16′

2. Construct a perpendicular from the principal point to this line and extend it through the curve.
3. Construct the tangent to the point of intersection of the perpendicular and the curve. This is the trace of the apparent horizon.

B. Swing of oblique photograph
1. Draw a line passing through the principal point parallel to the horizon trace; this line is the principal parallel.
2. The swing is measured as the angle between the principal parallel and the fiducial x axis of the oblique. If the forward end of the principal parallel is above the x axis, record the swing as plus (+), if below, minus (−).

C. Depression angle of oblique photograph
1. The dip angle between the true horizon plane and the apparent horizon plane is a function of the altitude of the exposure station above the earth.

Dip angle (in minutes)

$$= .98\sqrt{H \text{ (in feet)}}. \qquad (18.1)$$

where H is the altitude in feet above mean terrain and the dip angle is in minutes of arc. The relationship between the dip angle and altitude is shown as table 18-1. The formula given above is derived in section 18.5.1.1.
2. Draw the principal line through the principal point perpendicular to the horizon trace.
3. Measure the distance along the principal line from the principal point to the apparent horizon and divide by the focal length of the camera. This gives the tangent of the apparent depression angle—the angle between the camera axis and the plane of the apparent horizon.
4. The depression angle θ is obtained as the sum of the dip angle and the apparent depression angle.

D. y tilt of the vertical photograph
1. The y tilt of the vertical photo is measured in the plane of the oblique as the angle between the incline and a line parallel to the horizon.

$$t_{yo} = S - \Psi. \qquad (18.2)$$

where S is the swing of the oblique and Ψ is the relative swing of the oblique. S and Ψ must be used with the proper sign. If t_{yo} is plus (+), the forward end of the vertical photo (i.e., the positive print) is tipped down from the horizontal; if minus (−), up.
2. Project the oblique y tilt to a vertical plane by the formula:

$$\tan t_y = \tan t_{yo} \cos \theta. \qquad (18.3)$$

where θ is the depression angle of the oblique photo, and t_y is the y tilt of the vertical photo measured in a vertical plane perpendicular to the principal plane of the oblique photo.
3. Since t_{yo} is usually small and θ is close to 30°, the approximate formula:

$$t_y = 0.87\, t_{yo}. \qquad (18.3a)$$

may be used without serious error.

E. x tilt of the vertical photo
Obtain the x tilt (t_x) of the vertical photograph as the difference between the depression angle and the locking depression angle of one of the corresponding oblique photographs.

$$t_x = D - \theta. \qquad (18.4)$$

If the sign of t_x is plus (+), the side of the vertical photo toward the oblique photo which was used is tilted up from the horizontal; if minus (−), down.

F. Break angle
1. Obtain the angle Z between the principal plane and the vertical plane perpendicular to the isoline for each oblique by the formula:

$$\tan Z = \tan t_{yo} \sin \theta. \qquad (18.5)$$

The sign of angle Z is the same as the sign of angle t_{yo} of oblique y tilt.
2. The total break angle B between the principal planes of the two obliques is the algebraic sum of the relative setting and the two Z angles.

$$B = Z_{90} + Z_{45} + RS. \qquad (18.6)$$

OFEK07700

When the depression angles are close to 30°, $\sin\theta = 0.5$ approximately, and the break angle may be found by the approximate equation:

$$B = t_{90} + .5S. \qquad (18.6a)$$

3. If the principal plane of the left oblique is taken as the base, a plus (+) sign indicates B is measured in a clockwise direction; a minus (−) sign, counterclockwise.

18.4 Radial Triangulation

The map positions of the trimetrogon photographs are located by means of radial triangulation using overlapping vertical and oblique photographs. The nadir of each trimetrogon set serves as a station from which directions to points imaged on the photographs are traced. These traces are recorded as paper templets from which metal templets are made.

18.4.1 RECTIFICATION OF RADIALS

The radials from oblique trimetrogon photographs must first be rectified. Figure 18-4 shows the oblique photograph and its relationship to the principal and horizon planes. O is the perspective center, P is the principal point, N is the nadir point, I is the isocenter, HX is the true horizon trace, and θ is the depression angle. Any image point A on the oblique determines a vertical plane passing through that point and the nadir. The angle between this plane $ANOX$ and the principal plane, measured in the plane of the oblique, is λ; measured in the horizontal plane it is ϕ. The relation between the angles ϕ and λ may be derived as follows:

$$\frac{\tan\phi}{\tan\lambda} = \frac{HX}{HO} \div \frac{HX}{NH} = \frac{NH}{HO} = \csc\theta$$

$$\tan\phi = \tan\lambda \csc\theta. \qquad (18.7)$$

Figure 18-5 shows a front and a side view of the principal plane OHN. In the side view, $P'I$ is parallel to OH, and OP' equals OP equals f, and

$$IH = OP' \sec\theta = OP \sec\theta = OB.$$

Therefore, if the oblique photo is rotated about the horizon trace into the horizon

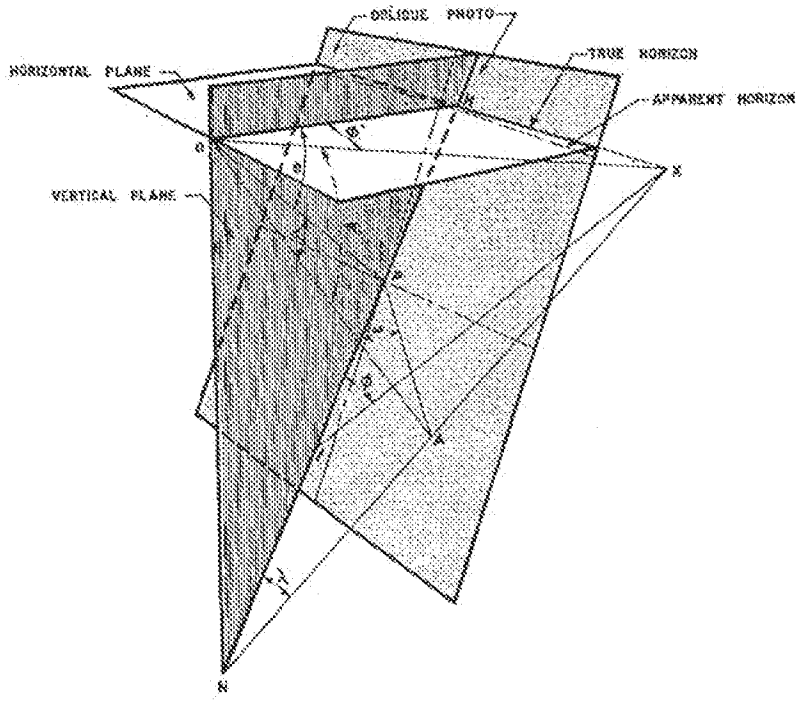


FIGURE 18-4. Geometric relations of a single oblique photograph.

OFEK07701

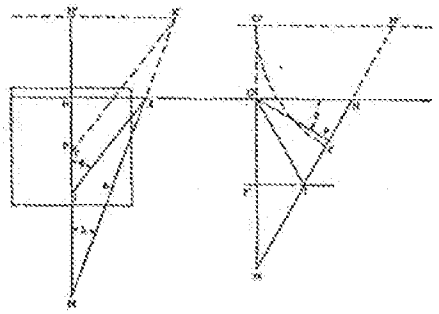


FIGURE 18-5. Theory of rectification of radials.

plane, the point I will coincide with the exposure station O, and the angle HIX is equal to the true horizontal angle ϕ. In other words, the true horizontal angle between nadir point radials is equal to the angle between the isocenter radials to the horizon intercepts of the nadir point radials. The instrument most generally used for the mechanical rectification of radials is the rectoblique plotter (fig. 18-6).

In figure 18-5 the dashed lines $H'O'$ and $H'X'$ represent the traces of an auxiliary plane parallel to the horizon plane. If an auxiliary ray center I' is chosen so that $H'I'$ equals $H'O'$ equals $H'N \sin \theta$, then

$$\tan H'I'X' = \frac{H'X'}{H'I'} = \frac{H'X'}{H'N} \sin \theta$$

$$= \tan \lambda \csc \theta = \tan \phi$$

and

$$H'I'X' = \phi.$$

18.4.2 PAPER TEMPLETS

Paper templets for each trimetrogon set are made by first marking a center point on the templet, representing the nadir point. The templet, on some suitable translucent and stable paper, is placed over the vertical photo with the center point of the templet and the nadir of the vertical photo coinciding. Rays are drawn from the center (nadir) to all the points previously picked. The rays from the corresponding oblique photos are added by means of the rectoblique plotter in the following manner:

A. Place one of the oblique photos in the rectoblique plotter with the horizon parallel to the horizon bar and set the photo and templet arms for the correct focal length and depression angles. Then orient the photo arm through the center pass point (annex point).
B. Place the paper templet nadir point on the templet arm pivot and orient the templet so that the templet arm passes through the corresponding annex point and tape the templet in place.
C. Check the radials to the side pass points against the photo positions. If necessary, revise the machine settings slightly to obtain coincidence.
D. Draw radials through all pass points, tie points, detail points, and control points, and mark them by appropriate symbol or number. If the detail points are to be intersected on the manuscript with the photoangulator, they are omitted from the paper templets.
E. Rotate the templet approximately 180° (180° plus or minus the break angle) and repeat steps A to D using the opposite oblique photograph. The use of the annex point for orientation eliminates the necessity for setting off the break angle.

The paper templets are then combined into a preliminary plot on a light-table. The nadir point spacing of the first two templets is determined by multiplying the photo base by the ratio of the manuscript and photo scales. After the templets from the photos of one flight strip have been combined, the points of intersection of the rays are marked on each templet. These marks indicate the proper length for the arms of the metal templets which will be made from the paper templets.

18.4.3 METAL TEMPLETS

Metal templets have generally been used in trimetrogon compilation because they are better adapted to the long narrow shape of the trimetrogon templet than are plastic or fiberboard sheets in which the rays have been punched. The paper templet is placed on a suitable flat surface and the rays of the metal templet are constructed by direct overlay and are bolted together. The templets and arms are marked for identification. The rays to the many detail points are not normally incorporated into the metal templets because the individual metal templets would become too cumbersome and because interference between the many arms would usually result.

18.4.4 MANUSCRIPT BASE

Templet plots for trimetrogon compilations normally cover an extensive area requiring the use of a large table or group of tables. The tables should be leveled and

OFEK07702

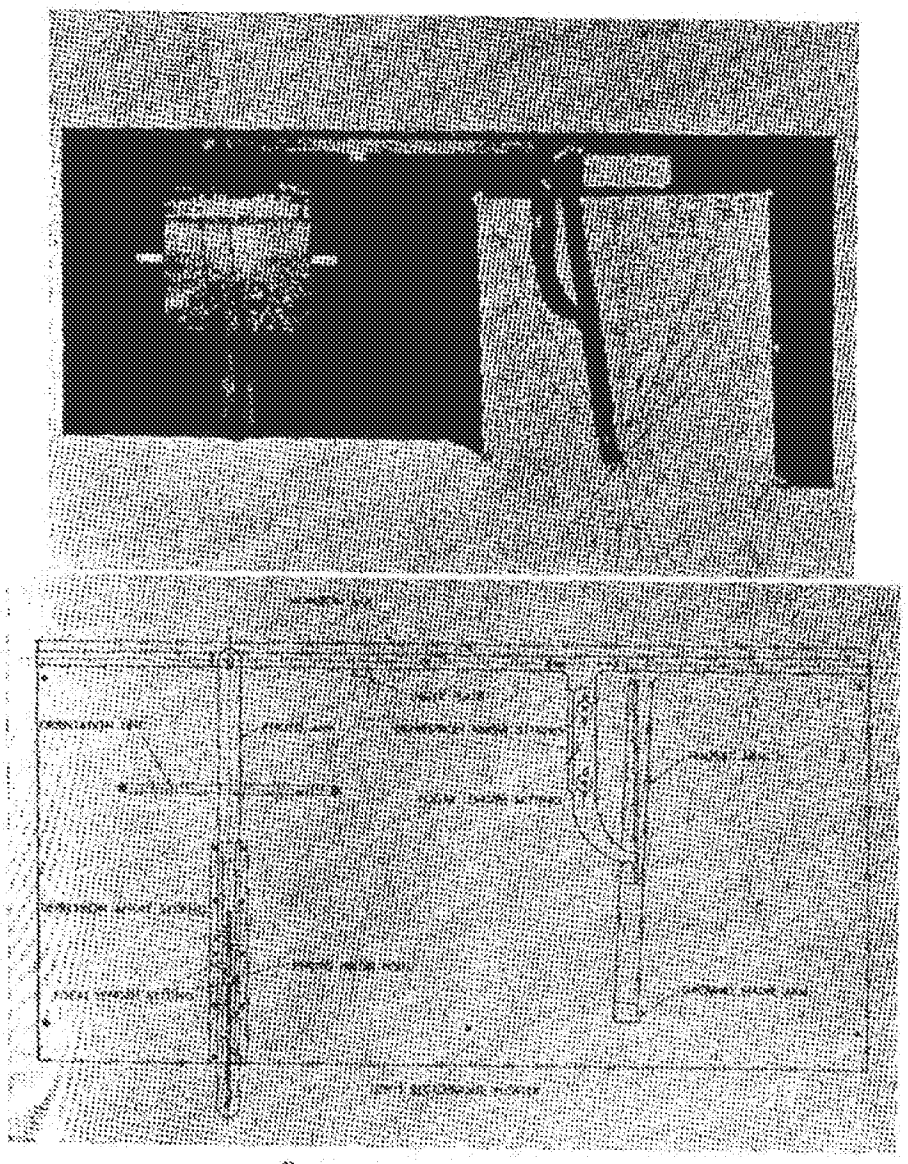


FIGURE 18-6. Rectoblique plotter.

should be constructed of material soft enough to take pins easily. A suitable manuscript base can be made of a stable plastic sheet such as Vinylite. A map projection is drawn on this base at the desired compilation scale (preferably at one-half the mean scale of the vertical photographs). The geographical control points are plotted on the manuscript and studs are pinned at each of these positions. Colored tabs should be placed over these studs to differentiate control points of varying reliability. The assembly of metal templets is then begun by first laying the templets which contain control points. The rays of the templet are laid over the studs of the corresponding control points. Studs are also placed at all other intersections (fig. 18–7). After all of the templets have been assembled, all ray intersections and

OFEK07703

nadir points are marked on the manuscript by circling and labeling pin holes made by lightly tapping pins inserted in each stud. The resulting product is the manuscript base. The initial paper templets are resected to the base and all of the detail points are added to complete the preparation of the point base. The photoangulator, a simple angle-measuring and drawing device, may be used to plot detail points on the manuscript base without the use of paper templets. With this device, the rays are drawn directly on the base from the photograph.

18.5 Delineation and Sketching

The planimetric detail, cultural features, hydrography, ridge lines, etc., which are to be portrayed on the manuscript should normally first be delineated on the photographs in brilliant opaque colors with the aid of a suitable stereoscope. This is done to facilitate the viewing of the photographic features and the manuscript base simultaneously, as required with sketchmaster-type compiling instruments. Even when using sketching equipment which does not firmly require delineation of the photographs, this technique still provides a convenient check system.

18.5.1 SCALE VARIATIONS

There are two distinct types of scale variation inherent in an oblique photograph. The scale along the horizon parallels (x scale) is constant for any one parallel although each successive parallel has a different scale. The scale along the principal line (y scale), varies constantly from point to point. The x and y scales at any image point on the oblique photograph can be found by the following equations:

$$S_x = \frac{d \cos \theta}{H} \tag{18.8}$$

$$S_y = \frac{d^2 \cos^2 \theta}{fH} \tag{18.9}$$

where d is the distance measured from the image point perpendicularly to the trace of the true horizon, θ is the depression angle, f is the principal distance, and H is the altitude of the exposure station above the ground point, d, f, and H being in the same units. Because the depression angle of trimetrogon obliques is usually about 30°, the scales at the principal point are:

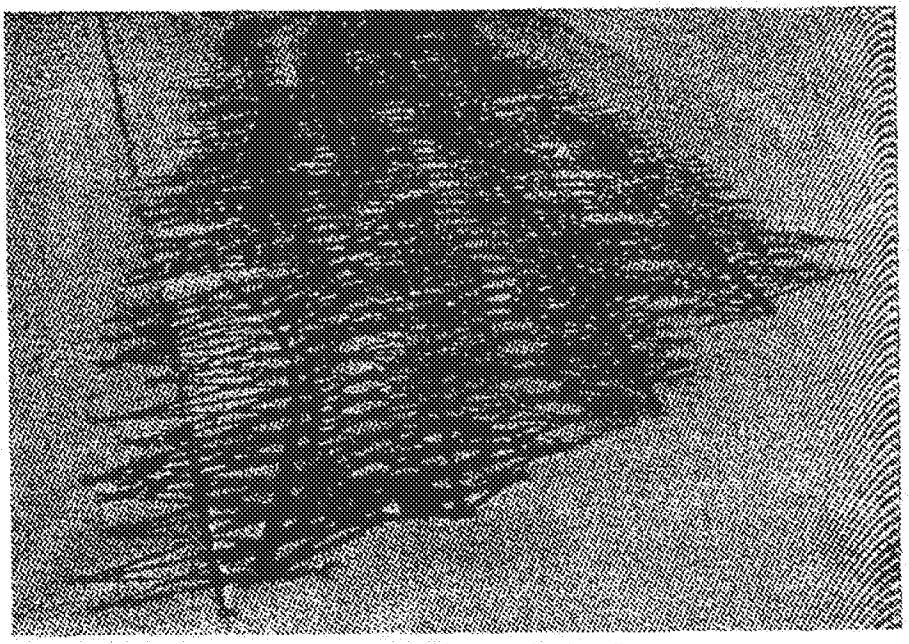


FIGURE 18-7. Metal templet assembly.

OFEK07704

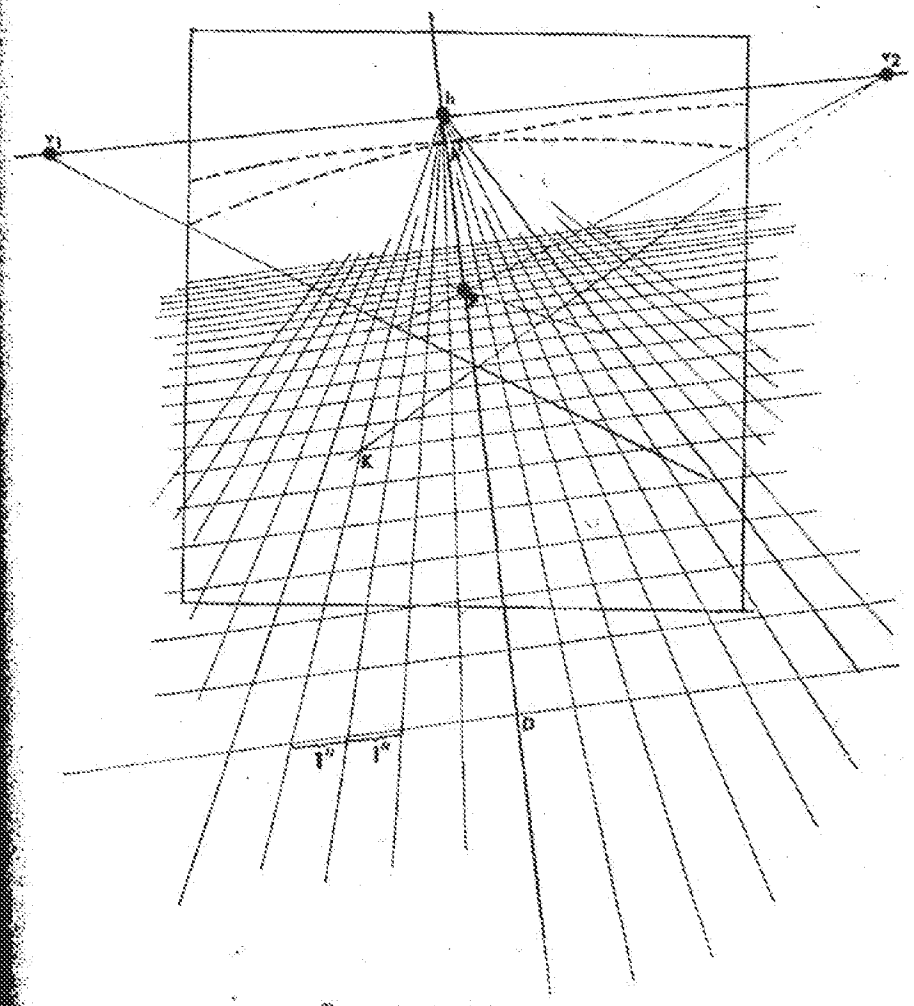


FIGURE 18-8. Perspective grid.

$$S_x = \frac{f}{2H} \qquad (18.8a)$$

$$S_y = \frac{f}{4H}. \qquad (18.9a)$$

18.5.1.1 PERSPECTIVE GRID

The use of perspective (Canadian) grids primarily suited to country having little lief. To use them, it is necessary that the pression angles and heights of the exsure stations be determined. Generally grid used on the photograph is a onent perspective quadrilateral type. The used in compilation of maps or charts of a squared character, the size of the ares being dependent on the scale at which the compilation is desired. The first step in preparing the grid is to determine and mark the principal point P of the photograph, and to delineate the apparent horizon (fig. 18–8). Then draw the principal line through P at right angles to the apparent horizon. Calculate θ', the depression angle of the optical axis from the apparent horizon, from

$$\theta' = \arctan \frac{PH}{f}.$$

Compute the vertical angle α (see fig. 18–9), subtended at the exposure station by rays to the apparent and true horizons, by use of the equation:

OFEK07705

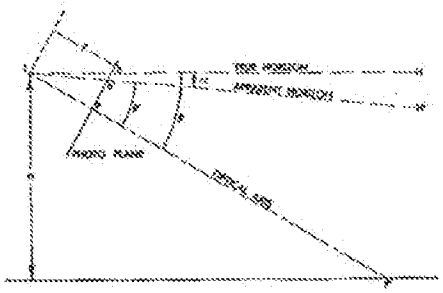


FIGURE 18-9. Angular relations of an oblique photograph.

$$\alpha = \text{Arc tan } 0.9216 \sqrt{\frac{2H}{R} + \frac{H^2}{R^2}}$$

in which H is the flight height above mean terrain, R is the radius of the earth, and the value 0.9216 is a constant to correct for atmospheric refraction. H and R must be in the same units. Where H is small in comparison with R, the term H^2/R^2 may be dropped without significant effect on the value of α, and the equation becomes:

$$\alpha \text{ (in minutes)} = .98\sqrt{H \text{ (in feet)}}$$

The depression angle θ of the optical axis from the true horizon is:

$$\theta = \theta' + \alpha.$$

Calculate hh', the distance between true and apparent horizons on the photograph:

$$hh' = f \tan \theta - Ph'.$$

Construct the true horizon through h perpendicular to the principal line (as in figure 18-8). Select a grid interval m (in feet). Establish the point D (as in figure 18-8):

$$hD \text{ (in inches)} = \frac{H \sec \theta}{m}.$$

Through the point D draw a line parallel to the true horizon. On it mark off divisions of one inch starting at D. Join these points with h. The resultant fan is the projection of a series of lines m distance apart on the ground all parallel to the principal line hPD. All lines at 45 degrees to the principal line vanish at the same point. If we establish this point we can draw the projection of the diagonals and so construct the projection of the squares. The vanishing points, one to the left and one to the right of h, are located on the true horizon a distance hv from h.

$$hv_1 = hv_2 = f \sec \theta.$$

Draw a line from the point v_1 through P (fig. 18-8). A series of lines at right angles to the principal line through the points intersected by the line Pv_1 gives the required grid. A similar line drawn from v_2 through P provides a check on drafting. It is not necessary to carry the divisions of the line through D out to the extent to provide coverage over the entire photograph. The divisions for the upper part of the grid may be obtained by extending, on any parallel, the equal divisions obtained on it by rays from D. These additional points are joined to h in the same way as the divisions from D. All lines parallel to Pv_1 vanish at v_1. A diagonal may be drawn to v_1 from any corner of any grid square. Thus the line hv_2 will locate the parallels for the bottom of the picture. The selection of the point P for the first diagonal was arbitrary. Figure 18-10 illustrates a Canadian grid (with ground distances and scales indicated) for an oblique photograph with a focal length of 6 inches, tilt angle of 60 degrees, and no swing.

18.5.2 SKETCHING INSTRUMENTS

The sketching instruments most generally used to transfer detail from trimetrogon photographs are the vertical and oblique sketchmasters. Both are forms of the camera lucida and are designed specifically for trimetrogon charting. The displacements caused by relief necessitate constant adjustment of the sketchmasters—the greater the displacement, the greater the adjustment. The amount of adjustment necessary is indicated by the detail points. Some mapping agencies prefer to utilize the stereoblique plotter instead of the oblique sketchmaster. The rectoplanigraph is a universal type which may be used for both verticals and obliques.

18.5.2.1 Electronic Sketchmaster

The U. S. Air Force has recently received a universal analog photographic rectification system which will permit the rapid transfer of detail from trimetrogon or any other type of aerial photographs, including panoramic. This instrument (fig. 18-11) is primarily the outgrowth of a prototype developed under the direction of the U. S. Naval Photo Interpretation

OFEK07706

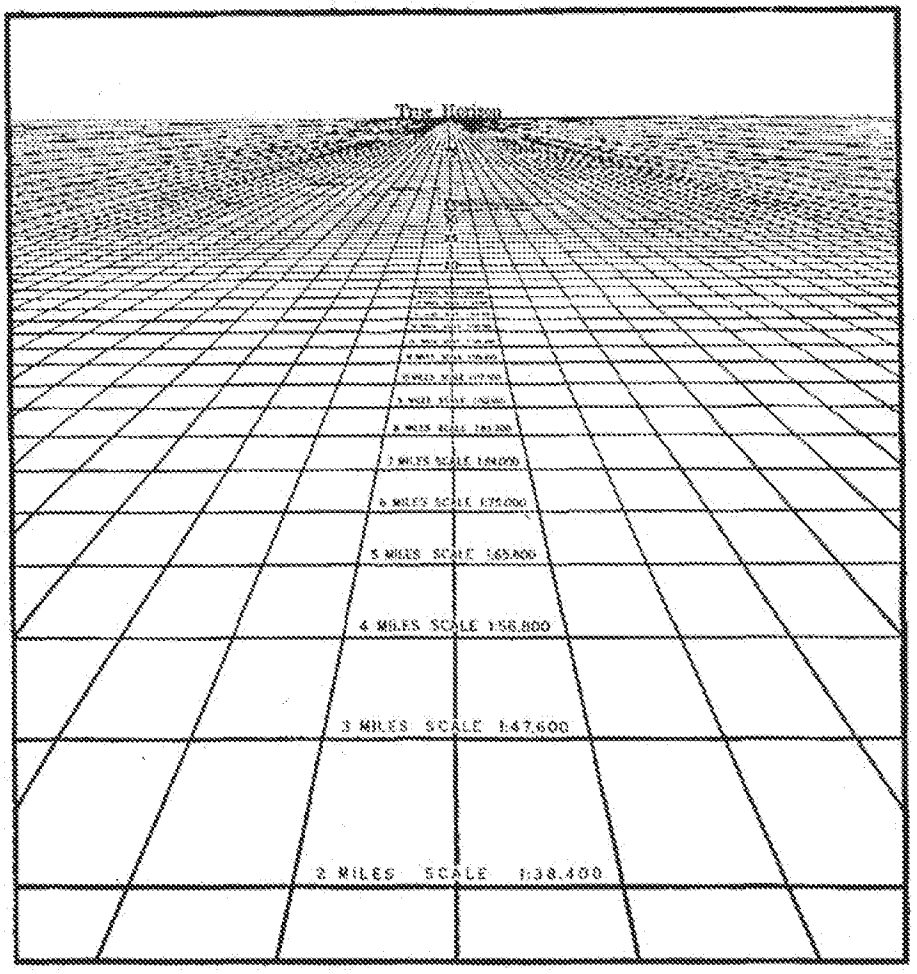


FIGURE 18-10. Canadian grid.

Center. The device consists of four basic components:

Input scanner—A lightweight portable coordinatograph which can be attached to any flat work surface.

Computer—An analog computer contained in the base of the plotter. The computer is programmed to accomplish plane and panoramic rectification, both corrected for earth curvature.

Console—A small desk-top unit containing a digital voltage readout and controls for adjusting, computing, and nulling the computer as well as scale factor and parallax controls for the plotter.

X–Y plotter—This is the output device consisting of a standard 30- by 30-inch plotter. It is desirable, for convenience of operation, to have the plotter operate in the vertical or near-vertical position.

The rectification equations are applications of the standard tilted-plane rectification method in the case of oblique photographs and of the cylinder-to-plane rectification in the case of panoramic photography. Corrections for relief displacement are not provided. The anticipated overall accuracy of the instrument will probably be in the neighborhood of 0.1 to 0.2 percent. The orienting of a photograph to a base can be accomplished by setting according to predetermined tilt and focal length or by empirical methods.

OFEK07707

INTRODUCTORY
Digital Image Processing



JOHN R. JENSEN

0876

OFEK06674

Introductory Digital Image Processing

OFEK06675

INTRODUCTORY DIGITAL IMAGE PROCESSING

A Remote Sensing Perspective

John R. Jensen

Department of Geography
University of South Carolina

PRENTICE-HALL, *Englewood Cliffs, New Jersey 07632*

0878

OFEK06676

Library of Congress Cataloging-in-Publication Data

Jensen, John R., (date)
Introductory digital image processing.

Includes bibliographies and index.
1. Remote sensing. 2. Image processing—Digital techniques. I. Title.
G70.4.J46 1986 621-36'78 85-19320
ISBN O-13-500828-X

Editorial/production supervision and
interior design: Arthur Hamparian
Cover design: Diane Saxe
Manufacturing Buyer: John Hall
Cover photo courtesy of COMTAL—3M, Inc.

© 1986 by Prentice-Hall
A Division of Simon & Schuster, Inc.
Englewood Cliffs, New Jersey 07632

All rights reserved. No part of this book may
be reproduced, in any form or by any means
without permission in writing from the publisher.

Printed in the United States of America

10 9 8 7 6 5 4

ISBN 0-13-500828-X 025

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Prentice-Hall of Southeast Asia Pte, Ltd., *Singapore*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*
Whitehall Books Limited, *Wellington, New Zealand*

OFEK06677

Contents

Preface *xi*

Acknowledgments *xiii*

1 *Introduction to Digital Image Processing* *1*

WHY COLLECT REMOTELY SENSED IMAGERY? 1
WHAT TYPE OF REMOTELY SENSED DATA SHOULD BE COLLECTED? 2
REMOTE SENSING RESOLUTION CONSIDERATIONS 4
WHY PROCESS THE REMOTELY SENSED DATA DIGITALLY? 5
EARTH RESOURCE ANALYSIS PERSPECTIVE 8
OVERVIEW OF IMAGE PROCESSING CONSIDERATIONS 8
REFERENCES 10

2 *Remote Sensing Data Acquisition Alternatives* *11*

INTRODUCTION 11
DENSITOMETER APPROACH 11
Optical-Mechanical Scanning 12
Video Digitization 16
Linear Photodiode or Charge-Coupled-Device Digitization 17

0880

OFEK06678

REMOTELY SENSED DATA ALREADY IN A DIGITAL FORMAT 19
The Landsat Multispectral Scanner and Thematic Mapper Sensor Systems 24
The SPOT Sensor System 36
Aircraft Multispectral Scanners 41
DIGITAL IMAGE DATA FORMATS 45
Band Sequential Format 47
Band Interleaved by Line Format 47
Band Interleaved by Pixel Format 47
Run-Length Encoding Format 47
REFERENCES 48

3 *Image Processing System Considerations* 50

INTRODUCTION 50
IMAGE PROCESSING SYSTEM CHARACTERISTICS 55
Number of Analysts and Mode of Operation 56
Central Processing Unit, Arithmetic Coprocessor, and Random Access Memory 57
Operating System and Compiler 58
Mass Storage 58
Display Resolution 59
Color Resolution 59
SOFTWARE 61
Introduction to the SENSOR Image Processing System 61
REFERENCES 61

4 *Initial Statistics Extraction* 63

INTRODUCTION 63
UNIVARIATE AND MULTIVARIATE STATISTICS COMPUTED FROM REMOTELY SENSED IMAGERY 63
THE HISTOGRAM AND ITS SIGNIFICANCE TO DIGITAL IMAGE PROCESSING OF REMOTE SENSOR DATA 69
REFERENCES 72

5 *Initial Display Alternatives* 73

INTRODUCTION 73
BLACK-AND-WHITE HARD-COPY IMAGE DISPLAY 74
Line Printer Brightness Maps 74

OFEK06679

Line Plotter Brightness Maps 77
Crossed-Line Shading 79
VIDEO IMAGE DISPLAY 82
Black-and-White and Color Brightness Maps 82
False Color Composites 89
TRANSFORMING VIDEO DISPLAYS TO HARD-COPY DISPLAYS 91
REFERENCES 92

6 *Image Preprocessing 94*

INTRODUCTION 94
RADIOMETRIC CORRECTION OF REMOTE SENSOR DATA 95
Detector Response Error 95
Atmospheric Effects 96
GEOMETRIC CORRECTION OF REMOTE SENSOR DATA 102
Geometric Rectification 104
REFERENCES 115

7 *Image Enhancement 117*

INTRODUCTION 117
IMAGE REDUCTION AND MAGNIFICATION 117
Image Reduction 117
Image Magnification 119
TRANSECTS 120
CONTRAST ENHANCEMENT 124
Linear Contrast Enhancement 125
Nonlinear Contrast Enhancement 129
RATIOING 135
SPATIAL FILTERING 138
Low-Frequency Filtering in the Spatial Domain 139
High-Frequency Filtering in the Spatial Domain 141
EDGE ENHANCEMENT 144
Linear Edge Enhancement 144
Nonlinear Edge Enhancement 149
SPECIAL TRANSFORMATIONS 151
Principal Components Analysis 151
Vegetation Indices 157
Texture Transformations 169
REFERENCES 173

OFEK06680

8 *Thematic Information Extraction* 177

INTRODUCTION 177
SUPERVISED CLASSIFICATION 178
The Classification Scheme 179
Training Site Selection and Statistics Extraction 186
Select the Appropriate Classification Algorithm 206
UNSUPERVISED CLASSIFICATION 215
Pass 1: Cluster Building 215
Pass 2: Assignment of Pixels to One of the Clusters Using Minimum-Distance Classification Logic 219
INCORPORATING ANCILLARY AND CONTEXTUAL DATA IN THE CLASSIFICATION PROCESS 222
LAND-USE CLASSIFICATION ACCURACY ASSESSMENT 225
Overall Classification Map Accuracy Assessment 225
Site-Specific Classification Map Accuracy Assessment 226
REFERENCES 230

9 *Change Detection* 234

INTRODUCTION 234
THE NATURE OF CHANGE DETECTION 234
CHANGE DETECTION ALGORITHMS 236
Image Differencing 237
Image Ratioing 241
Classification Comparisons 245
Preprocessing to Improve Change Detection 246
Change Vector Analysis 251
CONCLUSIONS 252
REFERENCES 252

10 *The Interface of Remote Sensing and Geographic Information Systems* 254

INTRODUCTION 254
FUNDAMENTAL GEOGRAPHIC INFORMATION SYSTEM CONCEPTS 255
Data Encoding 255
Data Management 262
Data Manipulation 263
Data Output 268
CONCLUSION 269
REFERENCES 269

OFEK06681

APPENDIX A: The SENSOR Digital Image Processing Programs 272

THE SENSOR PROGRAMS 272
Program Limitations 273
The Format of a SENSOR Image File 273
The STATISTICS File 273
AN OVERVIEW OF THE SENSOR SYSTEM PROGRAMS 274

APPENDIX B: Addresses of Public and Commercial Suppliers of Digital Image Processing Hardware and/or Software 364

Glossary 366

Index 375

OFEK06682

Preface

This text reviews the art and science of applying digital image processing techniques to remotely sensed imagery for the purpose of extracting useful earth resource information. Remotely sensed images of selected types of terrain are used to demonstrate the image processing topic under consideration. Whenever possible, the imagery is held constant, allowing the reader to compare the effect of applying various digital image processing techniques.

The level of the material presented assumes that the reader has completed either an airphoto interpretation or remote sensing course. The digital image processing algorithms are presented using mathematical notation. Whenever practical, the author works through the digital image processing manually to demonstrate how an algorithm is applied. The treatment assumes that the reader has a reasonable background in college algebra and univariate and multivariate statistics. The text can be used in an undergraduate or graduate one-semester course of introductory digital image processing, where the emphasis is on earth resource analysis. The digital image processing techniques presented are multidisciplinary in nature. They can be used in most earth science and social science research by scientists who believe that the analysis of remotely sensed data by computer may benefit their decision-making capability.

Listings of computer programs that perform many of the fundamental image processing functions are provided in Appendix A. The author assumes no liability for the accuracy of these programs, although great care was taken in their preparation. Please correspond with the author if you have suggestions concerning the image processing software. The programs are provided for the benefit of the reader, who may freely copy and use any and all parts of them. The author simply requests that their source be referenced when they are used for remote sensing research or educational purposes.

0885

OFEK06683

Acknowledgments

The author gratefully acknowledges the support and facilities of the Department of Geography at the University of South Carolina. Joseph Holmes of the department assisted in the design and drafting of the line drawings. Michael Hodgson assisted in the development and programming of the SENSOR digital image processing programs listed in Appendix A. Dr. Julian Minghi, chairman, provided resource support.

The American Society for Photogrammetry and Remote Sensing, the American Congress on Surveying and Mapping, the Association of American Geographers, and American Elsevier Publishing Company granted permission for the author to extract copyrighted material from his papers published in *Photogrammetric Engineering and Remote Sensing, The Manual of Remote Sensing, The American Cartographer, The Annals of the Association of American Geographers, The Professional Geographer* and *Remote Sensing of Environment.*

The high-resolution thermal infrared imagery of the Savannah River was obtained by the E. G. & G. Energy Measurements Group located in Las Vegas, Nevada.

The author appreciates the efforts of several anonymous reviewers, as well as those of Dr. John E. Estes and Earl Hajic of the University of California at Santa Barbara. Mr. Arthur Hamparian provided valuable technical and production assistance.

0886

OFEK06684

Introductory Digital Image Processing

OFEK06685

1

Introduction to Digital Image Processing

WHY COLLECT REMOTELY SENSED IMAGERY?

Earth scientists actively collect earth resource data to test hypotheses and simulate or model the environment. They collect the data using either in situ or remote sensing methods. One method of in situ data collection is to have a researcher located at the study area who is observing or questioning the phenomena under investigation. This often takes place in human geography and other social science research. In situ data collection may also be performed using transducers that are placed in direct contact with the phenomena of interest. Data recorded are normally analog electrical signals with voltage variations related to the physical variations measured. Often these data are transformed from analog signals into digital values using analog-to-digital (A-to-D) conversion procedures. Literally hundreds of types of transducers are available (e.g., thermometers, seismographs, and psychrometers). Measurements from these sensors provide much of the data for physical geography and other earth science research.

Remote sensing, however, departs from the aforementioned data collection methods because the sensor is remote from the phenomena; that is, it is not in direct physical contact with it (Colwell, 1984). Remote sensors usually record electromagnetic radiation (EMR) which travels at a velocity of 3×10^8 m s^{-1} from the source, directly through the vacuum of space or indirectly by reflection or reradiation to the sensor. As such, EMR represents a high-speed communications link between the sensor and remotely located phenomena. Changes in the amount and properties of the EMR become, upon detection, a valuable source of data for interpreting important properties of the phenomena with which they interact (Suits, 1983). Other types of force fields

0888

OFEK06686

may be used in place of EMR, including sound waves; however, the majority of remotely sensed data collected for geographic applications are the result of sensors that record electromagnetic energy (Jensen, 1983).

There are times when only in situ data collection is practical, as for example, in determining the socioeconomic characteristics of apartment households. There are other times when remote sensing is superior, as in the case of recording the spatial distribution of roof temperatures throughout a residential neighborhood. Often, however, it is necessary to collect both in situ and remotely sensed data. In such cases each type of data may be used to calibrate the other (Everett and Simonett, 1976).

WHAT TYPE OF DATA CAN BE REMOTELY SENSED?

To answer this question it is necessary to identify two classes of variables. First, there are variables that remote sensors can measure directly. This means that the remotely sensed data can provide fundamental biological and/or physical (*biophysical*) information directly, without having to use other surrogate or ancillary data. A good example is temperature mapping. A thermal infrared sensor can record the apparent temperature of a rock or water body by measuring the radiant flux emitted from the surface of the material. This is a true biophysical measurement. Such data are very useful in many physical science models. Another example is the determination of the precise x, y location or perhaps the height (z) of a particular object. Such information may be extracted directly from stereoscopic aerial photography and represents fundamental biophysical information. A list of nine selected biophysical variables that may be remotely sensed is found in Table 1–1. Jensen (1983) summarized the regions of the electromagnetic spectrum that are useful for measuring these variables when monitoring vegetation, soils and rocks, water, the atmosphere, and urban structure (Fig. 1–1).

The second general group of variables that may be remotely sensed are *hybrid* variables created by systematically analyzing more than one biophysical variable.

TABLE 1–1 Selected Fundamental Biophysical Variables Amenable to Remote Sensing

- Planimetric (x, y) location
- Topographic/bathymetric (z) elevation
- Object color
- Vegetation chlorophyll absorption characteristics
- Vegetation biomass
- Vegetation moisture content
- Soil moisture content
- Temperature
- Texture or surface roughness

Source: Jensen (1983).

OFEK06687

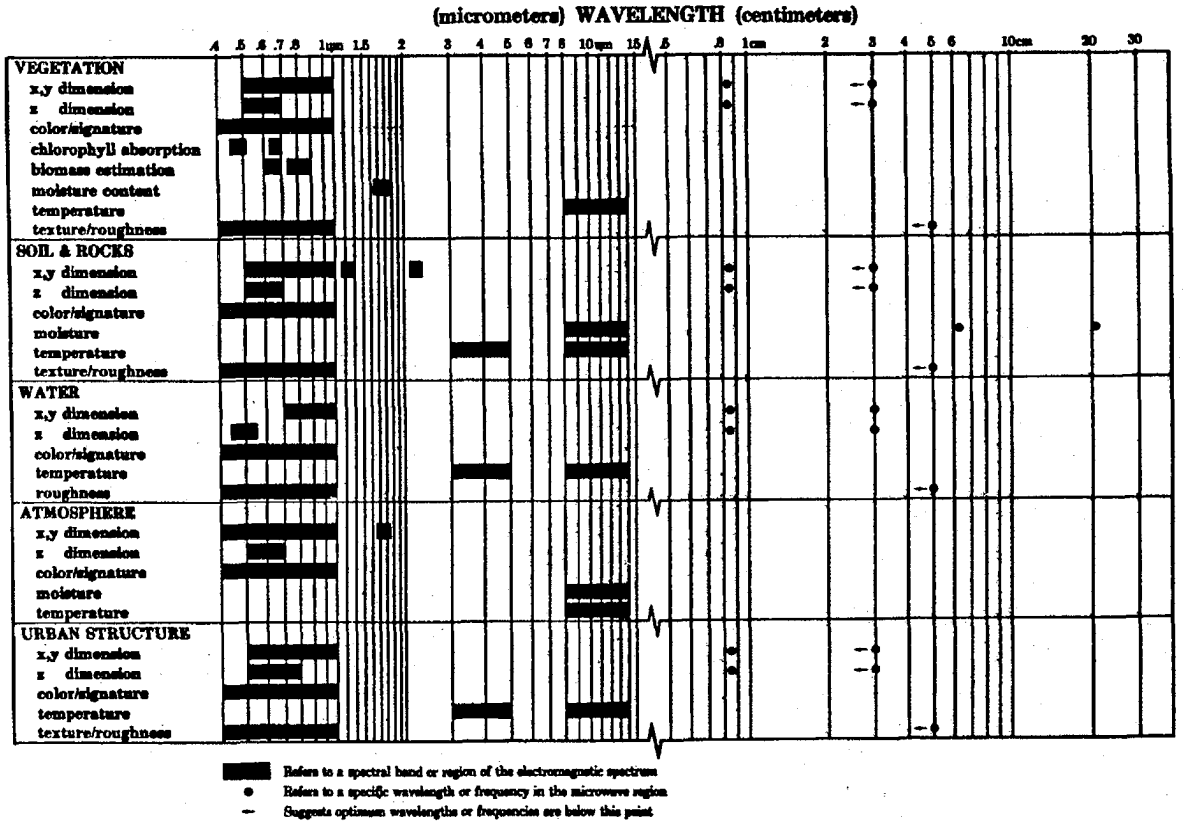


Figure 1–1 Regions of the electromagnetic spectrum useful for measuring biophysical variables associated with vegetation, soils and rocks, water, atmosphere, and urban structure. (From Jensen, 1983.)

For example, by remotely sensing a plant's chlorophyll absorption characteristics, its temperature, and moisture content, it may be possible to model these data to detect vegetation stress, a hybrid variable. The variety of hybrid variables is large; consequently, no attempt is made to identify them. However, it is important to point out that land-cover mapping represents a hybrid variable. The land cover of a particular point on an image is usually derived by evaluating several of the fundamental biophysical variables at one time [e.g., object color, location (x y), height (z), temperature, etc.]. So much attention has been placed on remotely sensing this hybrid *nominal*-scale variable that the fundamental biophysical variables that are usually *interval* or *ratio* scaled have been neglected. Certainly, nominal-scale land-use mapping is an important capability of remote sensing technology and should not be minimized. In fact, many social and physical scientists routinely use such data in their research. Quantitative remote sensing of the fundamental biophysical variables is still in its infancy. The extraction of biophysical information from remotely sensed data is destined to have a significant impact on most models that may incorporate such spatially distributed data (Estes et al., 1980).

The author believes that the majority of such biophysical information lies hidden

OFEK06688

or dormant within digital remotely sensed data, waiting to be extracted. To extract quantitative biophysical information from digital remote sensor data requires that the analyst understands (1) the types of resolution inherent in a given remote sensor system configuration, (2) the physics of how electromagnetic energy interacts with the phenomena of interest, (e.g., soils, vegetation, water, etc.), and (3) correct application of digital image processing techniques to extract biophysical or "hybrid" information. This book is concerned primarily with a description of how to apply digital image processing techniques. It assumes that the reader has a fundamental background in a science (e.g., geography, geology, forestry, etc.) to which the remote sensing technology may be applied. Prior to discussing the nature and application of digital image processing techniques, it is useful to describe several remote sensing resolution considerations so that a common level of understanding may be established.

REMOTE SENSING RESOLUTION CONSIDERATIONS

Resolution (or resolving power) is a measure of the ability of an optical system to distinguish between signals that are spatially near or spectrally similar (Swain and Davis, 1978). The ability to measure a biophysical variable using remote sensing requires careful consideration of four types of resolution: spectral, spatial, temporal, and radiometric resolution. The nature of each of these types of resolution must be understood by the scientist in order to extract meaningful biophysical information from the remotely sensed imagery.

Spectral resolution refers to the dimension and number of specific wavelength intervals in the electromagnetic spectrum to which a sensor is sensitive. The size of the interval or band may be large (i.e., coarse), as with panchromatic black-and-white aerial photography (0.4 to 0.7 μm) or relatively small (i.e., fine), as with band 3 of the Landsat 5 thematic mapper sensor system (0.63 to 0.69 μm). In the first instance, the sensor records all the reflected blue, green, and red radiant flux incident to it. In the second case, the sensor records only a very specific range of the red radiant flux. Certain regions or bands of the electromagnetic spectrum are optimum for obtaining information on biophysical parameters (Fig. 1–1). The bands are normally selected to maximize the contrast between the object of interest and its background (i.e., object-to-background contrast). Careful selection of the spectral bands may improve the probability that a feature will be detected, then identified and biophysical information extracted.

There is a relationship between the size of a feature to be identified and the spatial resolution of the remote sensing system. *Spatial resolution* is a measure of the smallest angular or linear separation between two objects that can be resolved by the sensor (Swain and Davis, 1978). Aerial photography spatial resolution is normally measured by engineers as the number of resolvable line pairs per millimeter on the image (Slater, 1980). For other sensor systems it is simply the dimension of the ground-projected instantaneous-field-of-view (IFOV) of the sensor system that moves across the terrain collecting information. For example, the Landsat multispec-

OFEK06689

tral scanner (MSS) ground spatial resolution is approximately 79 × 79 m. Generally, the better the spatial resolution, the greater the resolving power of the sensor system. Furthermore, a useful rule is that in order to detect a feature, the spatial resolution of the sensor system should be less than half the size of the feature measured in its smallest dimension. For example, if one wanted to identify the location of all maple trees within a city park, the minimum acceptable spatial resolution would probably be half the diameter of the smallest maple tree crown. Even this spatial resolution, however, will not guarantee success if there is no difference between the spectral response of the maple trees (the object) and the soil or grass surrounding them (the background).

Temporal resolution of a sensor system refers to how often a given sensor obtains imagery of a particular area. Ideally, the sensor obtains data repetitively to capture unique discriminating characteristics of the phenomena of interest. For example, agricultural crops typically have unique crop calendars in each geographic region (Jensen and Chery, 1980). Often in order to measure specific agricultural variables, it is necessary to acquire remotely sensed data at critical dates in the phenological cycle (Hanuschak, 1979). Analysis of multiple-date imagery provides information on how the variables are changing through time. Change information provides insight into processes influencing the development of the crop.

Radiometric resolution defines the sensitivity of a detector to differences in signal strength as it records the radiant flux reflected or emitted from the terrain. It defines the number of just-discriminable signal levels; consequently, it may be a significant element in the identification of scene objects. The multispectral scanner on Landsats 1 to 5 initially recorded the reflected radiant flux in 6 bits and then expanded the data in three of the bands to 7 bits (values ranging from 0 to 127). The spatial ground resolution was approximately 79 × 79 m. Radiant flux recorded by the thematic mapper sensor on Landsats 4 and 5 recorded data in 8 bits (values from 0 to 255) at 30 × 30 m spatial resolution in six of seven bands. Thus, the latter system has improved radiometric and spatial resolution.

It is generally believed that improvements in resolution increase the probability that phenomena may be remotely sensed more accurately (Everett and Simonett, 1976). The trade-off is that any improvement in resolution usually will require additional data-processing capability for either human or computer-assisted analysis.

WHY PROCESS THE REMOTELY SENSED DATA DIGITALLY?

Human beings are adept at visually interpreting images produced by certain remote sensing devices, especially cameras. One could ask: Why try to mimic or improve on this capability? First, there are certain thresholds beyond which the human interpreter cannot detect just-noticeable-differences in the imagery. For example, it is commonly known that an analyst can discriminate only about 8 to 16 shades of gray when interpreting continuous-tone black-and-white aerial photography. If the data were originally recorded with 256 shades of gray, there may be more subtle

OFEK06690

information present in the image than the interpreter can extract visually. Furthermore, the interpreter brings with him or her all the pressures of the day, making the interpretations generally unrepeatable. Conversely, the results obtained by computer are almost always repeatable (even when wrong!). Also, when it comes to keeping track of a great amount of detailed quantitative information such as the spectral characteristics of a vegetated field throughout a growing season for crop identification purposes, the computer is much more adept at storing and manipulating such tedious information and possibly making a more definitive conclusion as to what crop is being grown. This is not to say that digital image processing is superior to visual image analysis. This is certainly not the case. Rather, there may be times when a digital approach is better suited to the problem at hand.

But what about the actual processes of manual (visual) versus computer-assisted (usually digital) image analysis? Are there similarities between the goals and methods of both procedures? Estes et al. (1983) suggest that there exist several image analysis tasks and basic elements of image interpretation that the manual and computer-assisted approaches share (Fig. 1–2). First, both manual and digital analysis of remotely sensed data seek to detect and identify important phenomena in the scene. Once identified, the phenomena are usually measured and the information used in problem solving. Thus both manual and digital analysis have the same general goals. However, the attainment of these goals may follow significantly different paths. For example, most of the fundamental elements of image interpretation are used in manual image analysis (Avery and Berlin, 1985). These include tone/color (the most fundamental elements), size, shape, texture, pattern (all spatial arrangements of tone/color), height, shadow, site, and association. The human mind is quite adept at synthesizing these complex elements of image interpretation and bringing to bear ancillary information based on one's background to test various hypotheses.

Conversely, most of the computer-assisted image processing to date has involved the use of only a few of the basic elements of image interpretation. In fact, the overwhelming majority of all digital image analysis appears to be dependent primarily on just the tone/color of objects in the scene using fundamental statistical pattern recognition techniques. Various techniques have been used to measure and incorporate additional elements of image interpretation into the image analysis process. For example, numerous studies have synthesized "texture" information from the spectral data in the imagery. Also, several attempts at "contextual" classification make use of neighboring pixel values, thus incorporating some level of "association" information. It should be stressed, however, that even though the computer-assisted use of several of these elements of image interpretation has been demonstrated, the majority of the image processing is still based on the analysis of multispectral tone/color. Machines have difficulty dealing with size, shape, texture, pattern, height, and shadow, let alone the more complex elements of site and association. Some believe that the current research into symbolic reasoning and expert systems will allow the incorporation of most of the elements of image interpretation into computer-assisted image analysis (Estes et al., 1983).

This state of affairs suggests that computer-assisted image processing for earth

OFEK06691

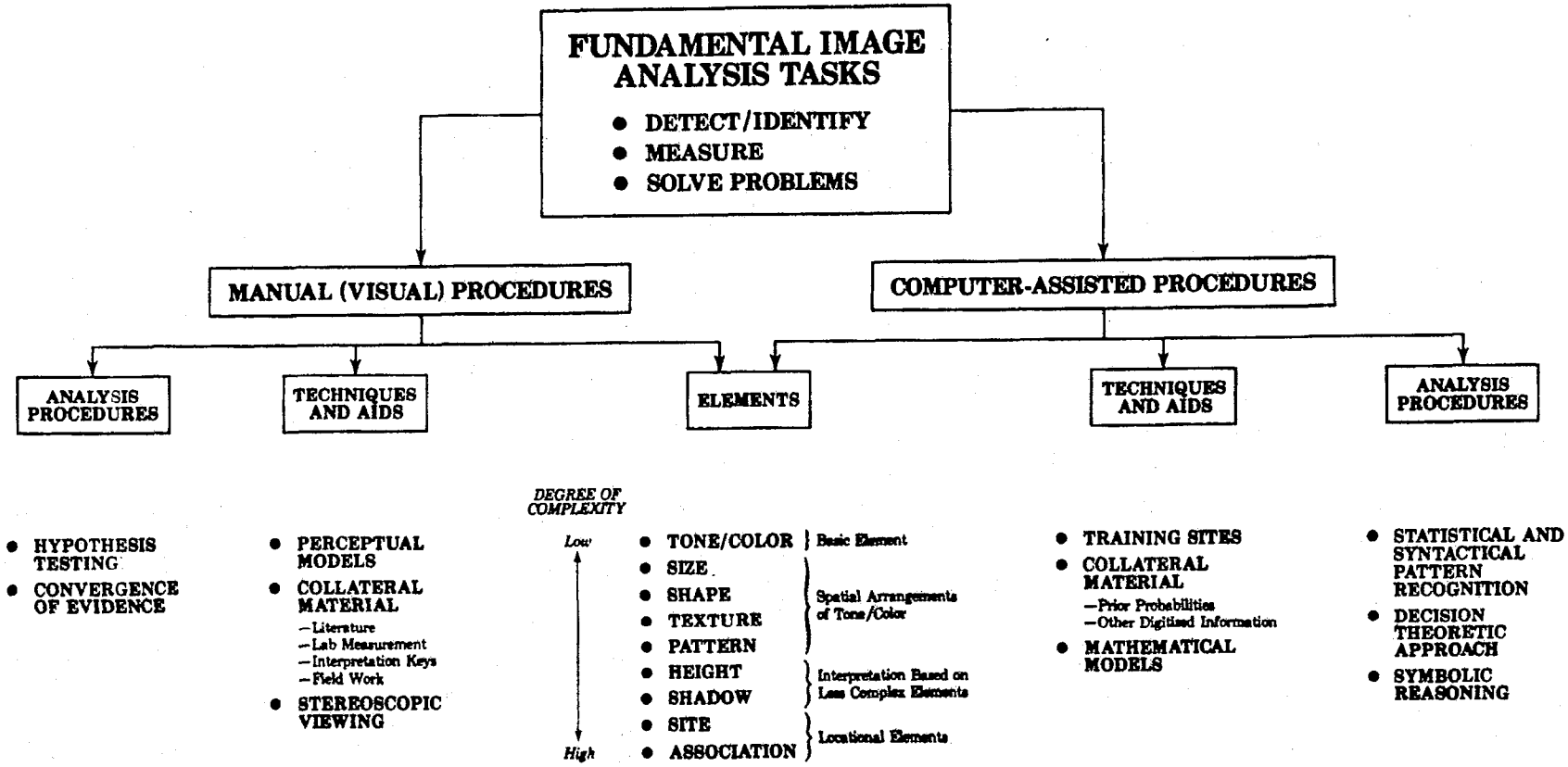


Figure 1–2 This conceptual diagram identifies human (visual) and computer-assisted image analysis procedures. Both procedures rely on analysis of the fundamental elements of image interpretation. Manual analysis at the present time incorporates many more of the complex elements in the analysis of remote sensing images. (After Estes et al., 1983.)

0894
OFEK06692

resource analysis is still in its infancy. If useful results can be extracted from imagery using even the most rudimentary image interpretation elements, there is hope that improved results will follow as more sophisticated elements are incorporated into the image analysis process.

EARTH RESOURCE ANALYSIS PERSPECTIVE

Geographers and other scientists are often concerned with the art and science of applying digital image processing techniques to remotely sensed data for the extraction of useful earth resource information. The mathematical bases of the techniques are important, but what is more important is how these techniques may be applied to real-world problems. Consequently, the digital image processing concepts and their mathematical underpinnings will be presented through the use of a few select digital remotely sensed data sets which are held constant throughout the book. These include:

1. A high-spatial-resolution (2.8 × 2.8 m) thermal infrared image (8.5 to 13.5 μm) of an area along the Savannah River in South Carolina obtained by an aircraft multispectral scanner on March 28, 1981, for some concepts involving single-image analysis
2. A seven-band Landsat satellite thematic mapper (TM) data set of Charleston, South Carolina, obtained on November 9, 1982, with a spatial resolution of 30 × 30 m for concepts involving multiple bands of remotely sensed data and for some single-image analysis
3. A Landsat multispectral scanner (MSS) data set of Denver, Colorado, obtained on September 30, 1976, and October 1, 1978, to demonstrate change detection algorithms at the urban fringe

In this manner readers will be able to determine for themselves whether an image enhancement or image analysis alternative provides useful information.

OVERVIEW OF IMAGE PROCESSING CONSIDERATIONS

The image analyst must evaluate several important image processing considerations prior to extracting meaningful information from remotely sensed data (Fig. 1–3). The analyst must first define the nature of the problem and identify all objectives and methods necessary to accept or reject research hypotheses. If a remote sensing approach to the problem is warranted, the analyst then evaluates several data acquisition alternatives. For example, he or she could digitize existing aerial photography or obtain the data already in a digital format from the EOSAT Company in Sioux Falls, South Dakota, where all Landsat MSS and TM data are archived. Then, assuming that the analysis is to be performed digitally, an appropriate digital image process-

OFEK06693

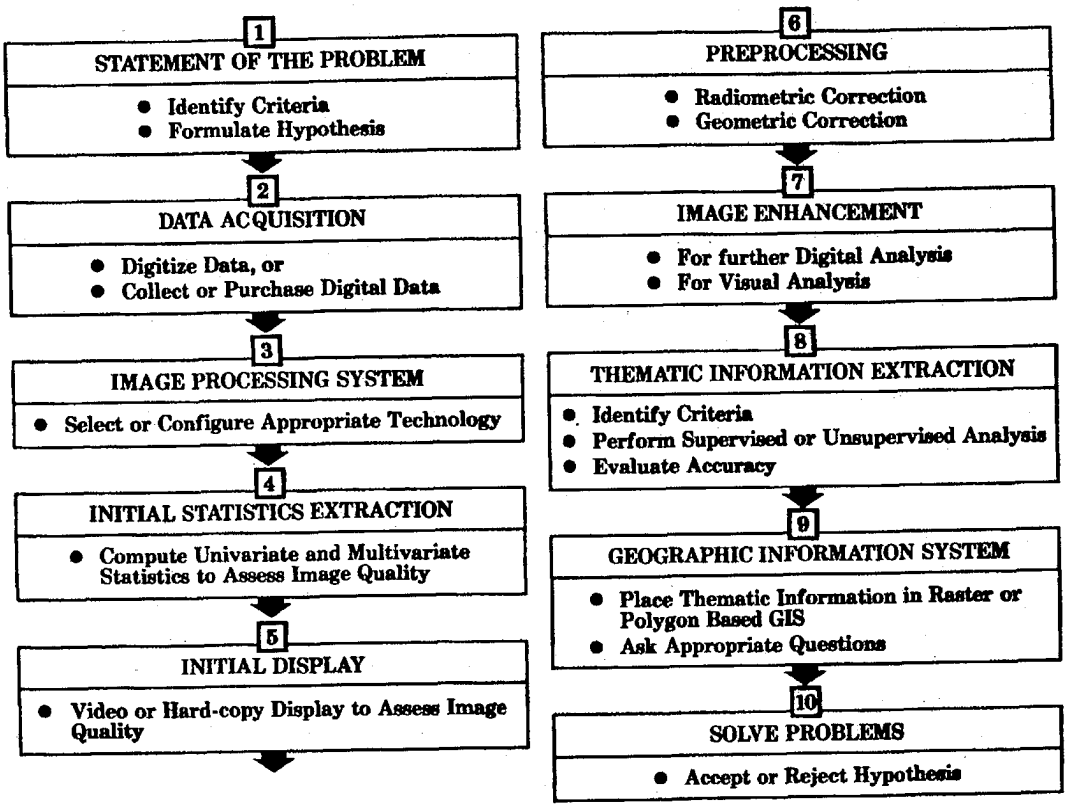


Figure 1–3 The analyst should evaluate these image processing considerations to extract meaningful information from remotely sensed data. The placement of the information in a geographic information system (GIS) is optional but often useful.

ing system is obtained. The analysis begins by first computing fundamental univariate and multivariate statistics used to display the raw digital remote sensor data. The statistics and initial display of the raw data provide valuable insight into image quality. Based on these observations the imagery is then preprocessed to reduce environmental and/or remote sensor system distortions in the data. This preprocessing usually includes both radiometric and geometric correction. Various image enhancements may then be applied to the corrected data for improved manual (visual) analysis or as input to further digital image processing. Important thematic information is then extracted from the imagery using either supervised (i.e., human assisted) or unsupervised techniques. Often, the thematic information gleaned from a supervised or unsupervised analysis is used by itself to solve problems. The remotely sensed thematic information may also be placed in a geographic information system (GIS) to improve its usefulness. A GIS can be used to relate the remote sensing–derived information with other types of spatially distributed data, such as soils, topography, and so on, to answer more complex problems. Each of the digital image processing considerations described above is addressed in the following chapters.

OFEK06694

REFERENCES

AVERY, T. E., and G. L. BERLIN. 1985. *Interpretation of Aerial Photographs*. Minneapolis, Minn.: Burgess Publishing Company, pp.27–28.

COLWELL, R. N. 1984. "From Photographic Interpretation to Remote Sensing," *Photogrammetric Engineering and Remote Sensing*, Vol. 50, pp. 1305–1307.

ESTES, J. E., J. R. JENSEN, and D. S. SIMONETT. 1980. "Impacts of Remote Sensing on U.S. Geography," *Remote Sensing of Environment*, Vol. 10, pp. 43–80.

ESTES, J. E., E. J. HAJIC, and L. R. TINNEY. 1983. "Fundamentals of Image Analysis: Analysis of Visible and Thermal Infrared Data," Chapter 24 in *The Manual of Remote Sensing*, R. N. Colwell, Ed., Falls Church, Va.: American Society of Photogrammetry and Remote Sensing, Vol. 1, pp. 987–989.

EVERETT, J., and D. S. SIMONETT. 1976. "Principles, Concepts, and Philosophical Problems in Remote Sensing," Chapter 3 in *Remote Sensing of the Environment*, J. Lintz and D. S. Simonett, Eds. Reading, Mass.: Addison-Wesley Publishing Company, Inc., pp. 85–127.

HANUSCHAK, G. 1979. *Obtaining Timely Crop Area Estimates Using Ground-Gathered and Landsat Data*. Washington, D.C.: United States Department of Agriculture Technical Bulletin, 1609.

JENSEN, J. R. 1983. "Biophysical Remote Sensing." *Annals of the Association of American Geographers*, Vol. 73, pp. 111–132.

JENSEN, J. R., and D. L. CHERY. 1980. "Landsat Crop Identification for Watershed Water Balance Determinations," *International Journal of Remote Sensing*, Vol. 1, pp. 345–359.

SLATER, P. N. 1980. *Remote Sensing: Optics and Optical Systems*. Reading, Mass.: Addison-Wesley Publishing Company, Inc.

SUITS, G. H. 1983. "The Nature of Electromagnetic Radiation," Chapter 2 in *The Manual of Remote Sensing*, R. N. Colwell, Ed. Falls Church, Va.: American Society of Photogrammetry and Remote Sensing, Vol. 1, pp. 37–60.

SWAIN, P. H., and S. M. DAVIS. 1978. *Remote Sensing: The Quantitative Approach*. New York: McGraw-Hill Book Company.

OFEK06695

2

Remote Sensing Data Acquisition Alternatives

INTRODUCTION

To perform digital image processing of remotely sensed data, it is necessary to have the imagery in a digital format. There exist two fundamental mechanisms for accomplishing this task:

1. Acquire remotely sensed imagery in analog format (e.g., an aerial photograph), then digitize it.
2. Acquire remotely sensed imagery in a digital format in the beginning, as with Landsat computer-compatible-tape (CCT) data.

It is instructive to review how data are placed in a digital format using these two alternatives.

DENSITOMETER APPROACH

Photographs are "hard-copy" pictorial records of a scene recorded by a camera. Photographs are recorded on photosensitive emulsions that may be either reflective, like a paper print, or transmissive, like a positive transparency (e.g., slide film). A photograph is two-dimensional and its reflectance or transmittance varies as a function of (x, y) position.

Aerial photographs are scaled-down images of objects that exist in reality. The gray shades or brightness values in such images may be evaluated at (x, y) positions

0898

OFEK06696

within the image and converted into discrete values that are proportional to the integrated output, reflectance, or transmittance of a small area usually called a picture element or *pixel*, centered on the (x, y) position (Fig. 2–1). The *density* of the (x, y) position is a measure of the light-absorbing capability of the silver or dye deposited at that location in the image.

Three major methods exist for converting existing hard-copy aerial photography, radar imagery, thermal infrared imagery, and so on, into a format suitable for digital image processing. These include: (1) optical-mechanical scanning, (2) video digitization, and (3) linear photodiode array or charge-coupled-device digitization.

Optical-Mechanical Scanning

A *densitometer* is a device that measures the average density of a small area of specified size on a photographic transparency or print. The measurement may be a meter reading or an electronic signal. When the area is smaller than a few hundred micrometers square, the instrument is called a *microdensitometer*. There are basically two types of scanning microdensitometers that are often used to convert hard-copy photographic images into digital values: flatbed and rotating drum systems. The imagery to be scanned (usually a film transparency) is placed on a flat glass surface when using a flatbed densitometer (Fig. 2–2). A single light source is then moved mechanically across the flat imagery in the x-direction. This light source is tracked by a photosensitive receiver (a photomultiplier tube) on the opposite side of the transpar-

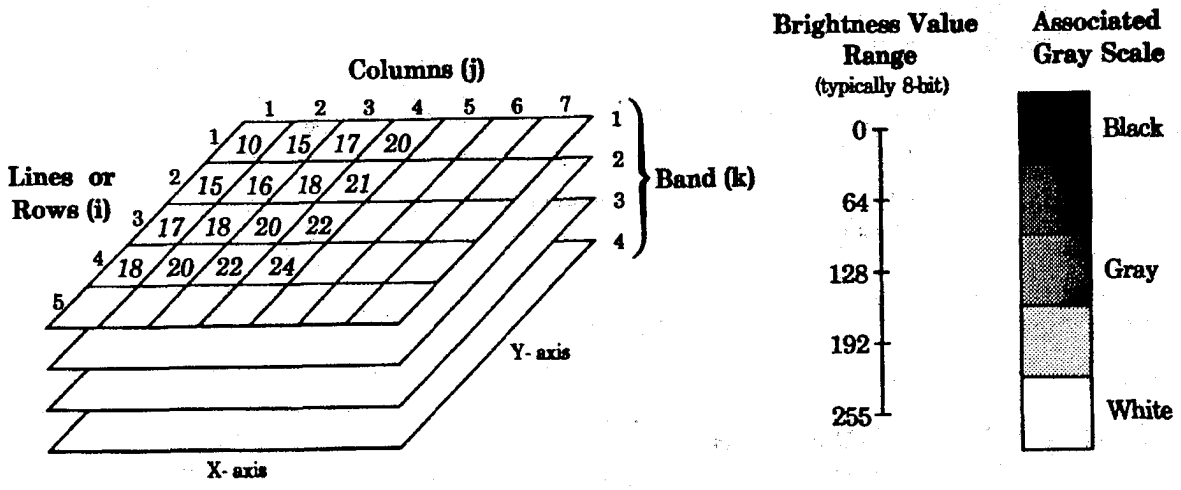


Figure 2–1 A digital remotely sensed image is typically composed of picture elements (pixels) located at the intersection of each row i and column j in each of the k bands of imagery. These data should be in perfect geometric registration. The brightness value (BV_{ijk}) at each pixel location is usually represented by a number ranging from 0 to 255. This value may be modulated to produce a gray shade on a CRT screen ranging from black (BV = 0) to bright white (BV = 255).

OFEK06697

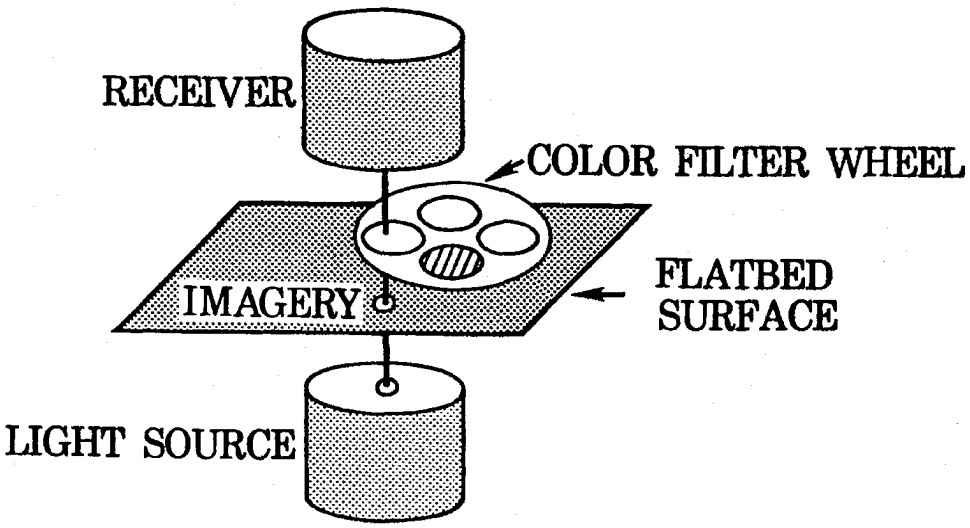


Figure 2–2 Fundamental elements of a flatbed scanning microdensitometer. The hard-copy remotely sensed imagery (usually a positive transparency) is placed on the flatbed surface. A very small light source (perhaps as small as 10 μm on a side) is moved mechanically across the flat imagery in the x-direction, emitting a constant amount of light. On the other side of the imagery, a receiver measures the amount of energy that passes through. When one line scan is complete, the light source and receiver step in the y direction some Δy to scan an area contiguous and parallel to the previous scan line. The amount of energy detected by the receiver along each scan line is eventually changed from an electrical signal into a digital value through an analog-to-digital (A-to-D) conversion. After the entire image has been scanned in this fashion, a matrix of brightness values is available for digital image processing purposes. A color filter wheel may be used if the imagery has multiple dye layers that must be digitizing. In this case the imagery is scanned three separate times using three different filters to separate it into its respective blue, green, and red components. The resulting three matrices should be in near-perfect registration representing a multispectral digital data set.

ency. At the end of each scan line, the light source steps in the y-direction some Δy, to scan along a line contiguous and parallel to the previous one. As the light source is scanned across the image, the continuous output from the receiver is converted to a series of discrete numerical values on a pixel-by-pixel basis. This analog-to-digital (A-to-D) conversion process results in a matrix (also called an array or raster) of values that are usually recorded as 8-bit values from 0 to 255 (Table 2–1 and Fig. 2–1). These data are then stored on a disk or tape for future analysis. Scanning imagery at spot sizes less than 15 μm may result in "noisy" digitized data because the spot size approaches the dimension of the film's silver halide crystals.

It is possible to obtain "multispectral" digital data by scanning color or color-infrared aerial photography three separate times using appropriate blue, green, and red filters held in a filter wheel (Fig. 2–2). This extracts spectral information from the respective dye layers found in color and color-infrared aerial photography and results in a registered, three-band digital data set for subsequent image processing.

Rotating drum optical-mechanical scanners digitize the imagery in a somewhat

OFEK06698

TABLE 2-1 Comparison of Image Digitizers

	Linear photodiode array	Vidicon	Rotating drum microdensitometer
Resolution	1728 × 2048	512 × 512	Selectable: 20,000 × 20,000 to 1250 × 1250
Format (scannable field)	No restrictions	No restrictions	9½ × 9½ in.
Maximum spatial Resolution at image plane			
a. 7.6 × 7.6 mm (¼ × ¼ in.) field	3.7 μm	12.4 μm	12.5 μm
b. 127 × 127 mm (5 × 5 in.) field	73.5 μm	248.0 μm	12.5 μm
c. 30 × 30 cm (12 × 12 in.) field	176.4 μm	595.3 μm	Field out of allowed range
Spatial precision	$\pm 10^{-5}$ pixel, along array axis	±2.5 pixel, pixel to pixel	±5 μm pixel to pixel
	$\pm 5 \times 10^{-2}$ pixel, scan line to scan line	±5 pixels, corner to corner	±2 μm rms over 1 cm
	±1 pixel, corner to corner, full frame		±25 μ, corner to corner
Gray-level precision	1 part per 1000 rms typical	1 part per 100 rms typical	2 parts in 256 rms typical
Data rate	200 kHz to MHz (dependent on bits per pixel)	16 MHz	30 to 60 kHz
Data depth	up to 12 bits/pixel	8 bits	8 bits
Image format	No size restrictions: transparencies reflective copy, real objects	No size restrictions: transparencies reflective copy, real objects	10 × 10 in. flexible transparencies or reflective copy
Time to scan full frame at maximum resolution	Variable 5 to 60 s	1/60 s	20,000 s (6 h)
Time to scan 5 × 5 in. area at maximum resolution	Variable 5 to 60 s at 74 μm pixel	1/60 s at 248 μm pixel	10,000 s at 12.5 μm pixel
Time to scan 5 × 5 in. area at 100 μm	Variable 3.7 to 44 s	N/A	159 s
Time to scan ½ × ½ in. area at full resolution	Variable 5 to 60 s at 7.4 μm pixel	1/60 s at 25 μm pixel	1016 s at 12.5 μm pixel
Power requirements	115 vac 15 A, 230 vac; 8-A line frequency; 48 to 65 Hz	115 vac 15 A, 60 Hz; 230 vac 8 A, 50 Hz	115 vac 15 A, 230 vac; 8-A line frequency; 48 to 65 Hz
Price	Low	Low	Medium–high

OFEK06699

different fashion (Figs. 2–3 and 2–4). The film transparency is mounted on a glass rotating drum so that it forms a portion of the drum's circumference. The light source is situated in the interior of the drum. The x-coordinate scanning motion is provided by the rotation of the drum. The y-coordinate is obtained by the incremental translation of the source–receiver optics after each drum revolution.

Optical-mechanical microdensitometers provide the highest spatial resolution (i.e., down to 1 μm) but are often slow (Table 2–1). For example, scanning a 5 × 5 in. area at 12.5-μm spatial resolution could require 10,000 s (2.76 h).

The National High Altitude Aerial Photography (NHAP) program is obtaining complete aerial photographic coverage of the conterminous United States from a high altitude to satisfy the needs of numerous federal and state agencies (U.S. Geological Survey, 1979). Stereoscopic aerial photography is acquired at 40,000 ft. using two aerial mapping cameras operating simultaneously; one with a 6-in. (152-mm) focal length lens with black-and-white panchromatic film, and another with an 8.25-in. (210-mm) lens and color-infrared (CIR) film. Thus panchromatic photography

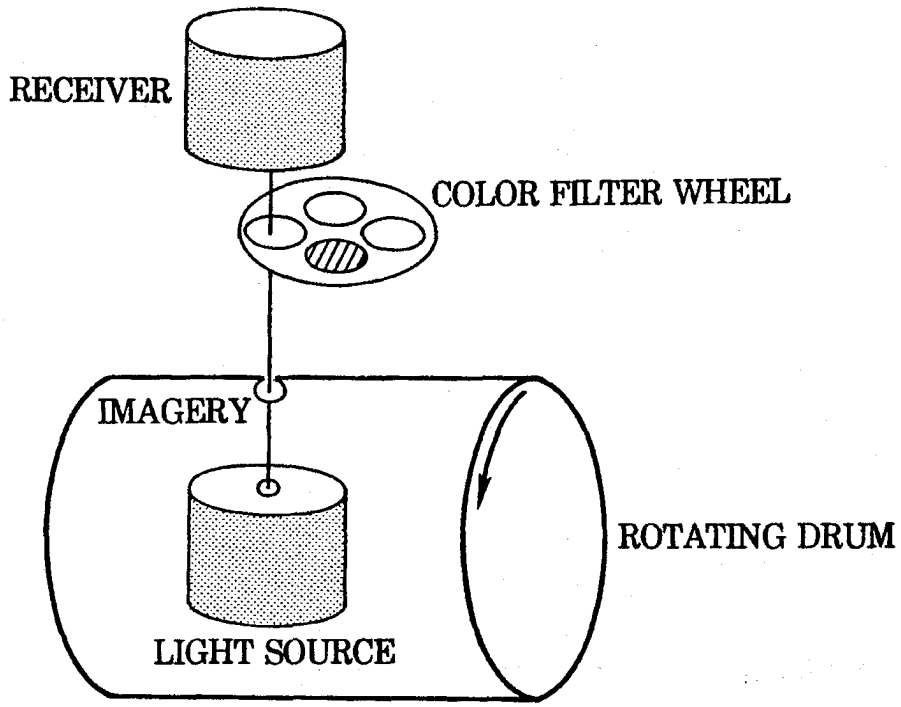


Figure 2–3 The rotating drum optical-mechanical scanning instrument works on exactly the same principle as the flatbed microdensitometer except that the remotely sensed data are mounted on a rotating drum so that it forms a portion of the drum's circumference. The light source is situated in the interior of the drum and the drum is continually rotated in the x-direction. The y-coordinate is obtained by the incremental translation of the source–receiver optics after each drum revolution. Some microdensitometers can write to film as well as digitize from film. In such cases the light source (usually a photodiode or laser) is modulated such that it exposes each picture element according to its brightness value. These are called "film writers" and provide excellent hard copy of remotely sensed data.

OFEK06700

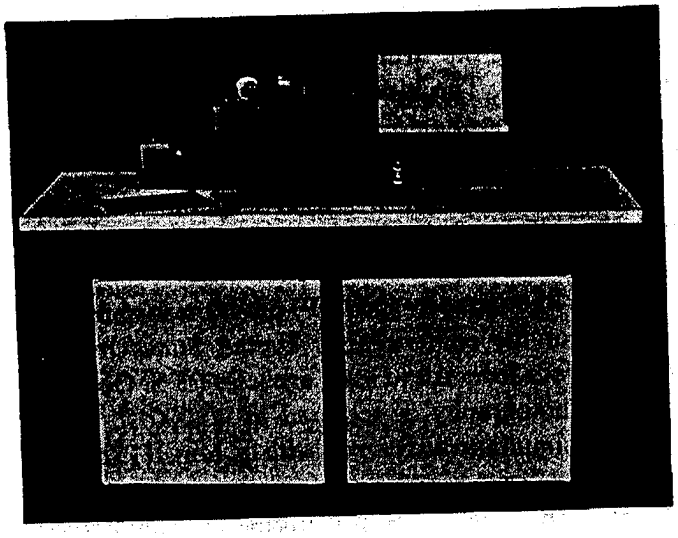


Figure 2–4 Optical-mechanical rotating drum digitizer and film writer. The photograph to be scanned is mounted on the drum so that it forms a portion of the drum's circumference. (Courtesy of Optronics International, Inc.)

is obtained at a 1:80,000 scale while CIR is obtained at a scale of 1:58,000. These images represent an important, extensive remote sensing data set useful for earth resource analysis. Holmquist et al. (1984) describe the photometric and geometric accuracies obtained when NHAP imagery is digitized using a 100-μm aperture for subsequent digital image processing.

Video Digitization

It is possible to digitize hard-copy imagery by sensing it through a video camera and then performing an analog-to-digital conversion on the 525 lines and 512 rows of data within the standard field of view (as established by the National Television System Committee). Video digitizing involves "freezing" and then digitizing a frame of analog video camera input. A full frame of video input can be read as rapidly as $\frac{1}{60}$ of a second (Table 2–1). A high-speed analog-to-digital converter, known as a *frame grabber*, digitizes the data and stores it in the buffer memory. The memory is then "read" by the host computer and the digital information is stored on disk or tape. A typical low-cost video frame-grabber configuration is shown in Fig. 2–5.

Video digitization of hard-copy imagery is performed very rapidly, but the results are not always useful for digital image processing purposes. For example, there are dramatic differences in the radiometric sensitivity and repeatability of various video cameras. One of the most serious problems is vignetting (or light falloff) away from the center of the image being digitized. This can affect the spectral signatures extracted from the scene. Also, any distortion in the vidicon optical system will be transferred to the digital remote sensor data, making it difficult to edge-match between adjacent images that have been digitized in this manner.

OFEK06701



Figure 2-5 Inexpensive video frame-grabber digitization system. The components include a black-and-white monitor, the vidicon camera, a filter wheel, a light table, and frame-grabber circuit board. This system can digitize a relatively small positive transparency image (usually an aerial photograph) into 240 × 256 pixels, with each pixel having a brightness value from 0 to 255 (an 8-bit range). The circuit board contains enough memory to hold the 240 × 256 8-bit brightness values until they can be transferred to more permanent storage on disk or tape. The black-and-white monitor is used to evaluate the nature of the image stored in the circuit board memory. (Courtesy of Spectral Data Corporation.)

Linear Photodiode or Charge-Coupled-Device Digitization

In the past, users of digital image processing equipment had to choose between the two digitizing technologies just discussed. Video digitizing represented a low-cost alternative which was convenient, but exhibited generally poor spatial resolution and often unrepeatable radiometric or gray-scale resolution. On the other hand, drum

OFEK06702

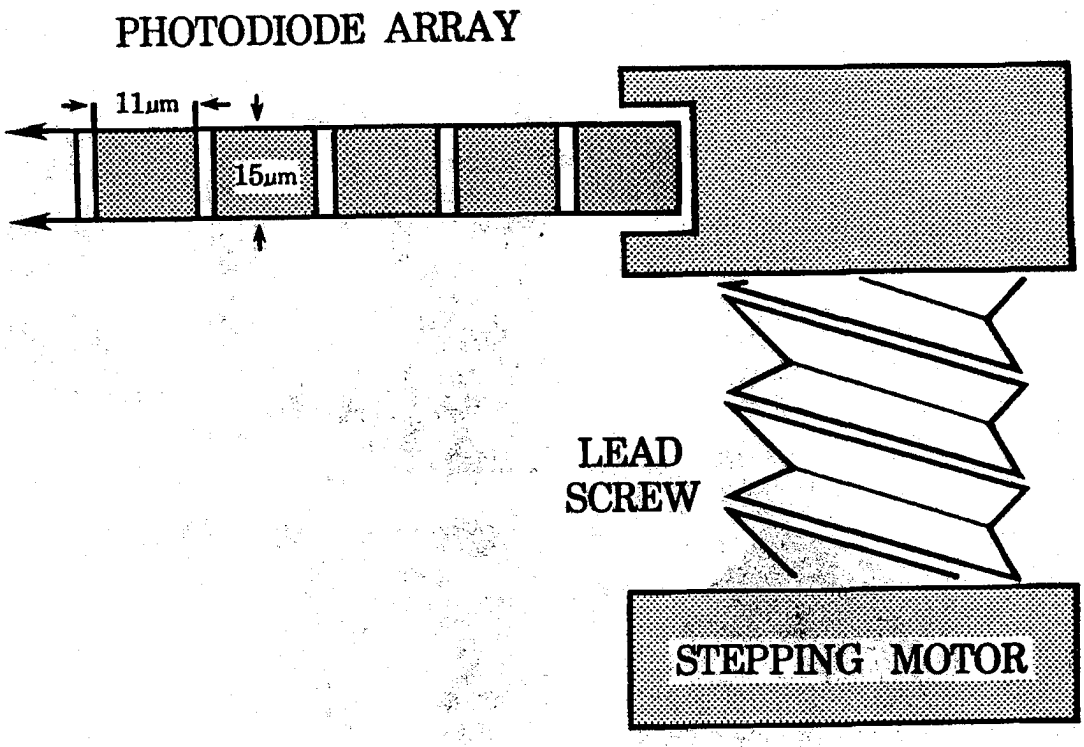


Figure 2-6 Logic of a linear photodiode array digitizer. The 1 × 2048 photodiode array is mounted on a pecision lead screw. Each of the 2048 detector elements in the photodiode array is fixed so that there is no movement in the x-direction. The image is scanned by moving the photodiode array in the y-direction. (Courtesy of Eikonix, Inc.)

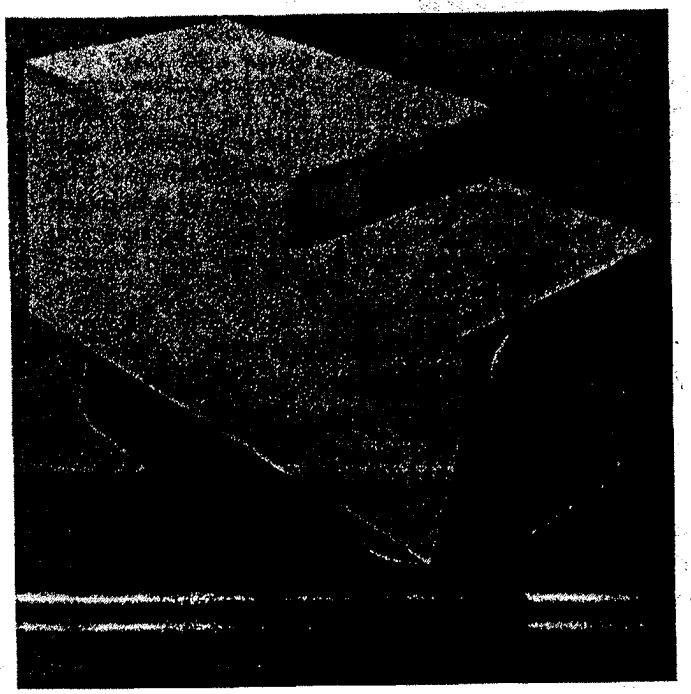


Figure 2-7 Camera head assembly that houses the linear photodiode array. (Courtesy of Eikonix Inc.)

OFEK06703

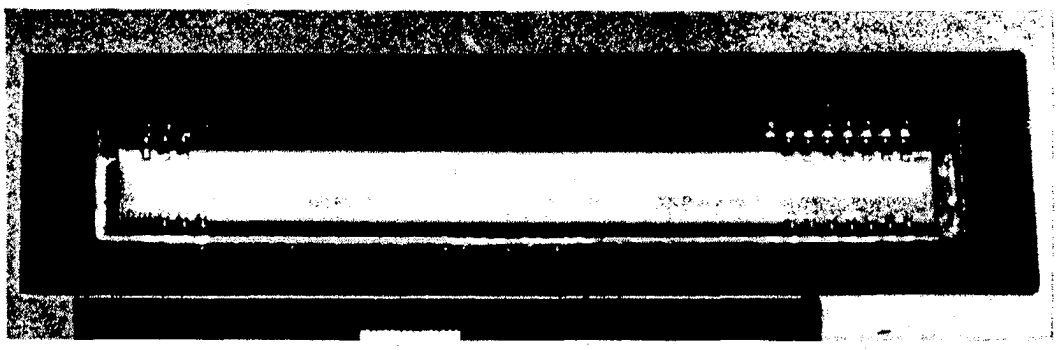


Figure 2–8 Line scan arrays are available as linear photodiode arrays or as charge-coupled devices (CCDs). This CCD array is composed of 1 × 2048 detector elements. (Courtesy of Eikonix Inc.)

and flatbed microdensitometers were slow, expensive, and difficult to maintain, but they offered repeatable spatial and radiometric accuracy. Recently, a new digitizing technology has been developed based on the use of photodiodes and/or charge-coupled devices (CCDs). In one such system, a linear photodiode array is mounted on a precision lead screw (Figs. 2–6 and 2–7). Because each detector element is fixed relative to all the other elements, and because the lens system is constant from scan to scan (assuming the same field of view), the spatial repeatability in the x-direction is more precise. The lead screw stepping motor can provide a positional repeatability of approximately ± 0.75 μm in the y-direction. Figure 2–8 is a photograph of a 2048-element charge-coupled device (CCD) that may be used instead of the photodiode array.

Typical output from such systems includes matrices of from 1728 × 2048 to 2048 × 3000 with 1 to 12 bits per pixel (values from 0 to 4096). Also, digitization generally takes place within 60 s. Thus, high spatial resolution, good radiometric accuracy, and fast digitization are provided. This technology is now a feasible alternative for converting hard-copy remotely sensed data into a digital format.

REMOTELY SENSED DATA ALREADY IN A DIGITAL FORMAT

Ideally, remote sensor data are provided to the user in a digital format, so no digitization is required. This is the case with much of the remotely sensed data now acquired. For example, Table 2–2 summarizes many of the remote sensor systems available for earth resource monitoring from 1980 to 1985 and their spatial, spectral, radiometric, and temporal resolutions. Many provide data in a digital format (Southward, 1985).

An overview of how remote sensing data are turned into useful information is shown in Fig. 2–9. First, the remote sensor system detects electromagnetic energy that exits from the phenomena of interest and passes through the sometimes murky atmosphere. The sensor system then transforms reflected or emitted electromagnetic energy that it has detected into an analog electrical signal which may subsequently

OFEK06704

TABLE 2-2 Remote Sensor Systems Functioning from January 1, 1980, to December 31, 1984, and Some of the Capabilities for Earth Resource Mapping[a]

Sensor	System resolutions: Spectral	Spatial (ground)	Temporal	Radiometric	Veg.	Soil–rock	Water	Atmos.	Urban
1 Landsat 3									
RBV—return beam vidicon									
Band	0.505–0.75μm	30 × 30m	18 days	6	A	G	L–N	Q	U–W
MSS—multispectral scanner									
Band 4	0.5–0.6	79 × 79	18 days	6	A	G	L,N	Q	U,V
Band 5	0.6–0.7	79 × 79	18 days	6	A,C	G	L	Q	U,V
Band 6	0.7–0.8	79 × 79	18 days	6	C	G	L–N	Q	U,V
Band 7	0.8–1.1	79 × 79	18 days	6	C	G	L–N	Q	U,V
Band 8	10.4–12.6	240 × 240	18 days	6	D	H	O	T	Y
2 Landsat 4, 5									
MSS	as above	as above	16 days	8	As above, except no band 8.				
TM—thematic mapper									
Band 1	0.45–0.52μm	30 × 30m	16 days	8	A,B	G	L–N	Q	U–W
Band 2	0.52–0.60	30 × 30	16 days	8	A	G	L,N	Q	U–W
Band 3	0.63–0.69	30 × 30	16 days	8	A,C	G	L	Q	U–W
Band 4	0.76–0.90	30 × 30	16 days	8	C	G	L–N	Q	U–W
Band 5	1.55–1.75	30 × 30	16 days	8	E	G	L	L	—
Band 6	10.40–12.5	120 × 120	16 days	8	D	H	O	T	U
Band 7	2.08–2.35	30 × 30	16 days	8	E	I	L	Q	U
3 NOAA Series									
AVHRR/2—advanced very high resolution radiometer									
Band 1	0.58–0.68	1.1 × 1.1k	14.5/day	8	A	G	O	Q	—
Band 2	0.725–1.10	1.1 × 1.1k	14.5/day	8	C	G	O	Q	—
Band 3	3.55–3.93	1.1 × 1.1k	14.5/day	8	—	—	O	R	—
Band 4	10.5–11.3	1.1 × 1.1k	14.5/day	8	D	H	O	Q,R	—
Band 5	11.5–12.5	1.1 × 1.1k	14.5/day	8	D	H	O	Q,R	—

0907

OFEK07705

4 SMS/GOES Series (East and West)									
VISSR—visible and infrared spin scan radiometer									
Visible	0.55–0.720	0.78 × 0.78k	0.5 h	6	A	G	L	Q	—
Infrared	10.5–12.6	7 × 7k	0.5 h	8	—	—	O	Q	—
VAS—Improved version of VISSR with 7 channels and 12 selectable filters. Can be operated in VISSR, multispectral imaging (MSI), or sounding mode.									
5 NIMBUS Series (Research and Development)									
CZCS—coastal zone color scanner									
Band 1	0.433–0.45	0.825 × 0.825k	6 days	8	A	G	L	Q	U
Band 2	0.510–0.53	0.825 × 0.825k	6 days	8	A	G	L	Q	U
Band 3	0.540–0.56	0.825 × 0.825k	6 days	8	P	G	P	Q	U
Band 4	0.660–0.68	0.825 × 0.825k	6 days	8	A	G	L	Q	U
Band 5	0.700–0.80	0.825 × 0.825k	6 days	8	C	G	L,M	Q	U
Band 6	10.5–12.5	0.825 × 0.825k	6 days	8	D	H	O	—	Y
6 HCMM—heat capacity mapping mission									
Band 1	0.5–1.10	500 × 500 m	Day	8	A,B	G	L,M	Q	—
Band 2	10.5–12.5	600 × 600	Day/night	8	D,E	H,J	O	T	—
7 SIR-A—Shuttle imaging radar (HH polarization; 50° incidence angle)									
L-band	23.5 cm	40 × 40 m	8 h	8	F	G,K	L,M	—	U,V
8 Aerial photography (color infrared)									
1: 10000	0.5–0.9	<0.3 m	Variable	Analog	A–C	G,K	L–N	Q	U–X,Z
1: 60000	0.5–0.9	3.0 m	Variable	Analog	A–C	G,K	L–N	Q	U–W,Z
1:130000	0.5–0.9	9.0 m	Variable	Analog	A–C	G,K	L–N	Q	U–W,Z
9 Aircraft multispectral scanner imagery (Daedalus DS-1260)									
Band 1	0.38–0.42	Variable	Variable	8	A	G	L,N	Q	U–X
Band 2	0.42–0.45	Variable	Variable	8	A	G	L	Q	U–X
Band 3	0.45–0.50	Variable	Variable	8	A,B	G	L	Q	U–X
Band 4	0.50–0.55	Variable	Variable	8	A	G	L	Q	U–X
Band 5	0.55–0.60	Variable	Variable	8	A	G	L	Q	U–X
Band 6	0.60–0.65	Variable	Variable	8	A	G	L	Q	U–X
Band 7	0.65–0.69	Variable	Variable	8	A	G	L	Q	U–X
Band 8	0.70–0.79	Variable	Variable	8	C	G	L,M	Q	U–X
Band 9	0.80–0.89	Variable	Variable	8	C	G	L,M	Q	U–X
Band 10	0.92–1.1	Variable	Variable	8	C	G	L,M	Q	U–X
Band 11	8.0–13.5	Variable	Variable	8	D,E	H	O	T	Y

0908

OFEK06706

TABLE 2-2 Cont.

Sensor	System resolutions: Spectral	Spatial (ground)	Temporal	Radiometric	Veg.	Soil–rock	Water	Atmos.	Urban.
10 TIMS—thermal infrared multispectral scanner									
Band 1	8.2–8.6	Variable	Variable	8	—	G–I	—	—	—
Band 2	8.6–9.0	Variable	Variable	8	—	G–I	—	—	—
Band 3	9.0–9.4	Variable	Variable	8	—	G–I	—	—	—
Band 4	9.4–10.2	Variable	Variable	8	—	G–I	—	—	—
Band 5	10.2–11.2	Variable	Variable	8	—	G–I	—	—	—
Band 6	11.2–12.2	Variable	Variable	8	—	G–I	—	—	—

[a] Earth resource mapping codes:

Vegetation
- A. Chlorophyll concentrations
- B. Carotenoid concentrations
- C. Biomass
- D. Surface temperature
- E. Moisture content
- F. Surface roughness

Soils and rocks
- G. Rock typing and lineament mapping
- H. Surface temperature (may include thermal inertia mapping)
- I. Hydrothermal alteration
- J. Moisture content
- K. Surface roughness

Water
- L. Areal extent
- M. Land/water boundary delineation
- N. Water penetration (may include turbidity)
- O. Surface temperature
- P. Gelbstoff (yellow stuff) concentration

Atmosphere
- Q. Daytime cloud cover mapping
- R. Nighttime cloud cover mapping
- S. Cloud/snow differentiation
- T. Cloud top temperature

Urban structure
- U. Anderson level I
- V. Anderson level II
- W. Anderson level III
- X. Anderson level IV
- Y. Surface temperature
- Z. Surface roughness

0909

OEEK0707

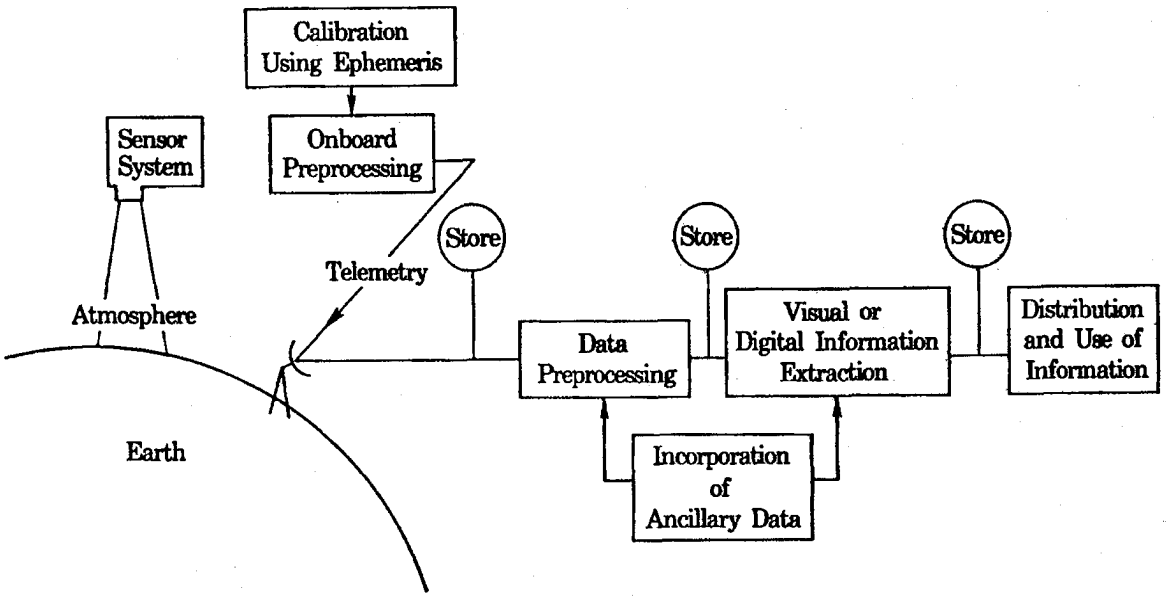


Figure 2-9 Overview of how remotely sensed data are turned into useful information. Once the remotely sensed data are collected, both on-board and ground preprocessing is performed to remove geometric and radiometric distortions. This may involve the use of ephemeris and/or ancillary data. The data are then ready for visual or digital analysis to extract meaningful information. This thematic information is then distributed and hopefully used. (After Landgrebe, 1983.)

be converted into a digital value. If an aircraft platform is used, the data are simply returned to earth. However, if a spacecraft platform is used, the data are usually telemetered to earth receiving stations. In either case, it may be necessary to perform some radiometric and/or geometric preprocessing of the remotely sensed data to improve its interpretability. Data may then be enhanced for subsequent visual analysis or they may be analyzed by computer using digital image processing algorithms. Information extracted from either approach is distributed and used to make decisions.

Throughout the process, remotely sensed data are stored in various analog and digital formats and often processed in conjunction with non-remote sensor data (i.e., ancillary data). Ancillary data might be used to rectify an image to a standard map projection, to improve its radiometric accuracy, or to stratify a study area prior to classification using information that can only be obtained on the ground (e.g., the subsurface geology of an area).

It is not possible to discuss all the sensor systems that can provide digital remote sensor data. However, it is possible to review four sensor systems that will continue to provide an extensive collection of remote sensor data of value for earth resource investigations. These include the Landsat multispectral scanner (MSS), the Landsat thematic mapper (TM) system, the SPOT HRV sensor, and a representative aircraft multispectral scanner. This discussion will provide insight into how digital remote sensor data are typically obtained, reformatted, and stored. Much of this

OFEK06708

book is concerned with the description of digital image processing techniques that can be applied to data obtained by these remote sensing systems.

The Landsat Multispectral Scanner and Thematic Mapper Sensor Systems

In 1967, the National Aeronautics and Space Administration (NASA) encouraged by the U.S. Department of the Interior, initiated the Earth Resource Technology Satellite (ERTS) program. This program resulted in the deployment of five satellites carrying a variety of remote sensing systems designed primarily to acquire earth resource information. The chronological launch and subsequent retirement history of some of the satellites are shown in Fig. 2–10. The ERTS-1 satellite launched on July 23, 1972, was the first experimental system designed to test the feasibility of collecting earth resource data by unmanned satellites. Prior to the launch of ERTS-B on January 22, 1975, NASA renamed the ERTS program *Landsat*, distinguishing it from the Seasat oceanographic satellite launched June 26, 1978. At this time ERTS-1 was retroactively named Landsat 1 and ERTS-B became Landsat 2 at launch. Landsat 3 was launched March 5, 1978; Landsat 4 on July 16, 1982, and Landsat 5 on March 1, 1984. A variety of mechanical failures prompted the retirement of some of the Landsat satellites. For a detailed account, refer to the *Landsat Data*

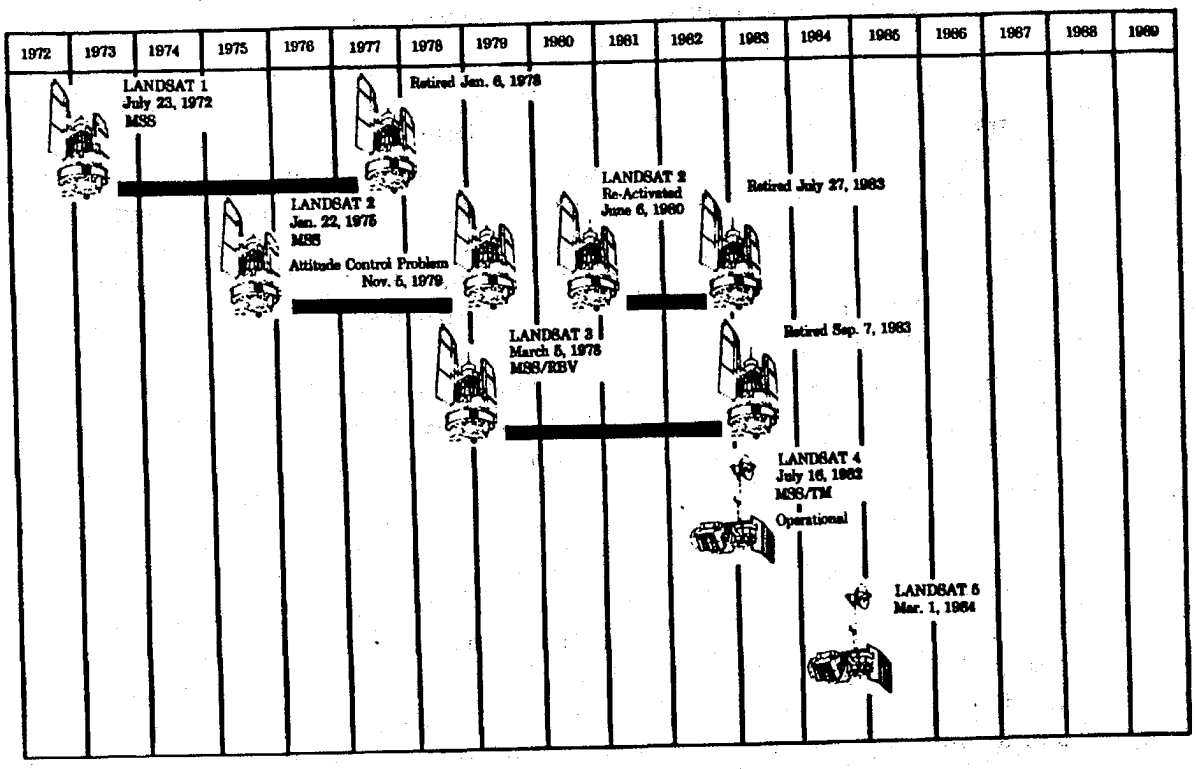


Figure 2–10 Chronological launch and retirement history of the Landsat series of satellites (1, 2, 3, 4, and 5) from 1972 to 1984.

OFEK06709

User Notes published by the EROS Data Center at Sioux Falls, South Dakota (NOAA, 1975–1985).

Orbital characteristics of Landsats. Landsats 1 to 3 were launched into circular orbits about the earth at a nominal altitude of approximately 919 km (570 mi). The sensor system platform is shown in Fig. 2–11. The satellites orbited the earth once every 103 min, resulting in 14 orbits per day (Fig. 2–12). The satellites had an orbital inclination of 99°, which made them nearly polar (Fig. 2–13). This caused them to cross the equator at an angle of approximately 9° from normal. This sun-synchronous orbit meant that the orbit plane precessed about the earth at the same angular rate at which the earth moved about the sun. This characteristic caused the satellites to cross the equator at approximately the same local time (9:30 to 10:00 A.M.) on the illuminated side of the earth.

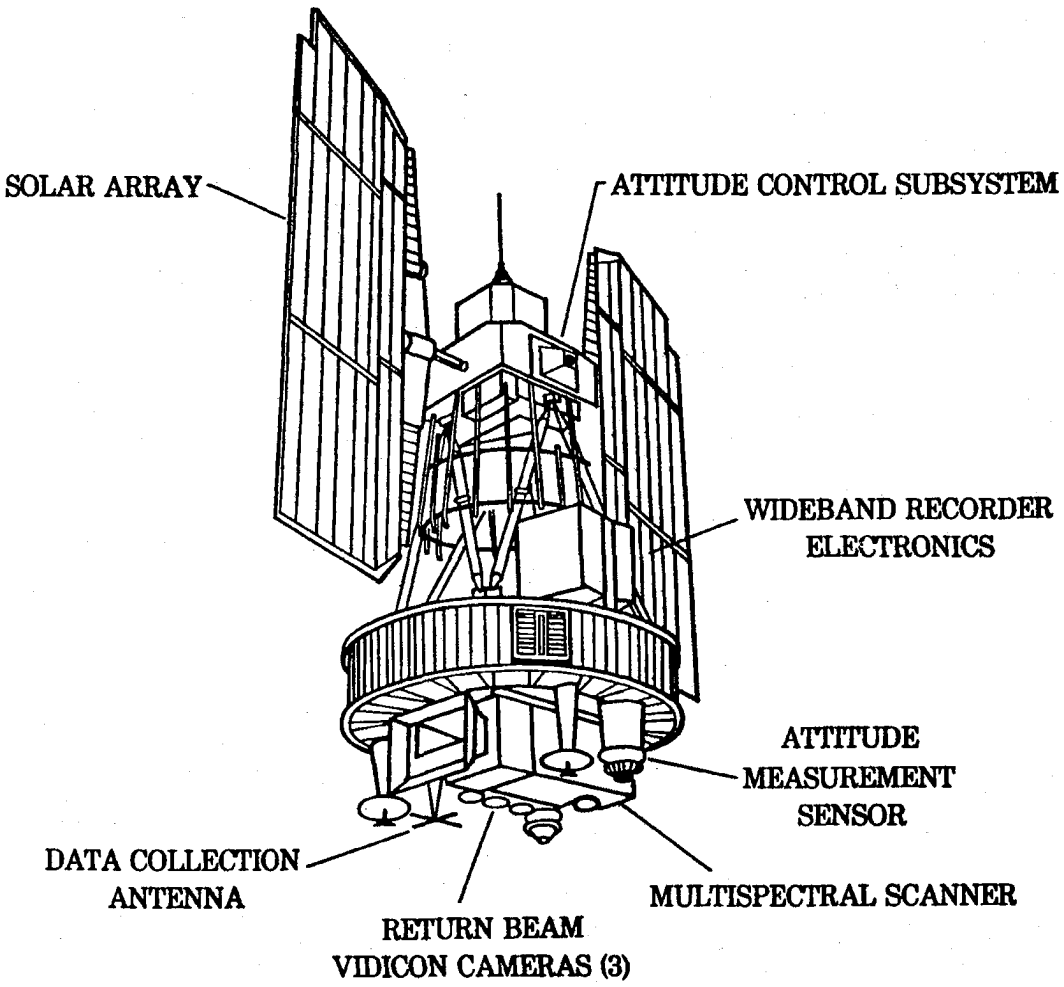


Figure 2-11 Nimbus-style platform used for Landsats 1, 2, and 3 and associated sensor and telecommunication systems. (From *NASA Landsat Data Users Handbook*.)

OFEK06710

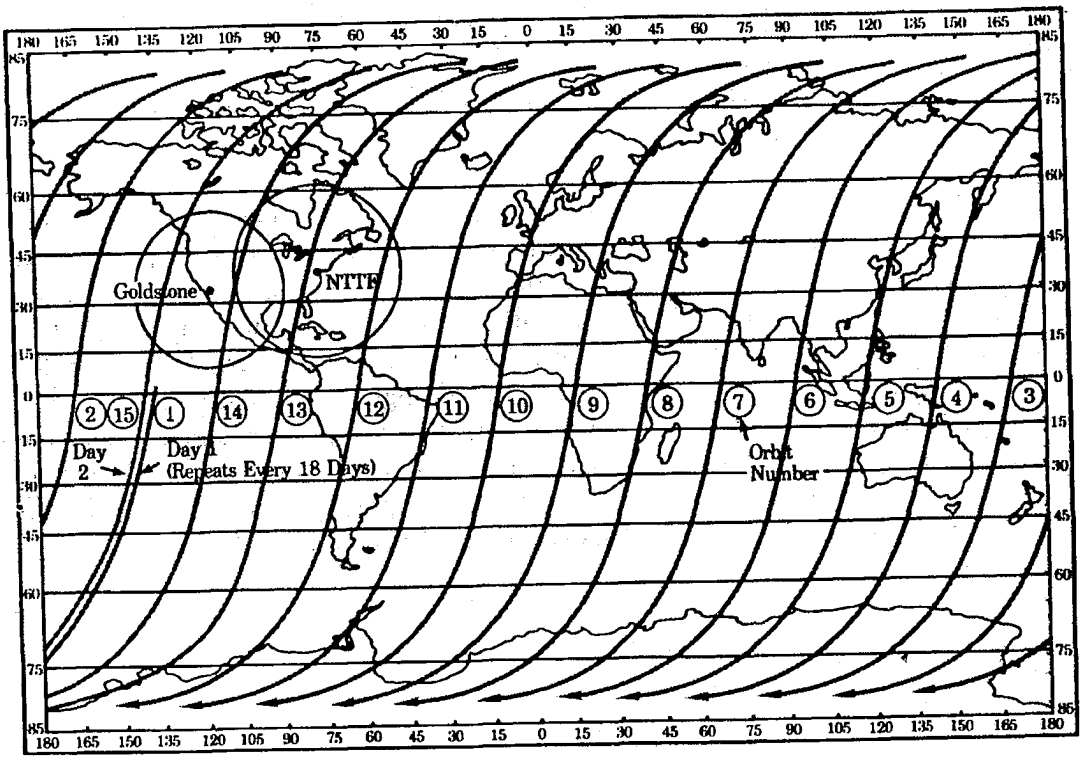


Figure 2–12 Orbital tracks of Landsat 1, 2, or 3 during a single day of coverage. The satellite crossed the equator every 103 min, during which time the earth rotated a distance of 2875 km under the satellite at the equator. Every 14 orbits 24 h elapsed. (From Taranik, 1978.)

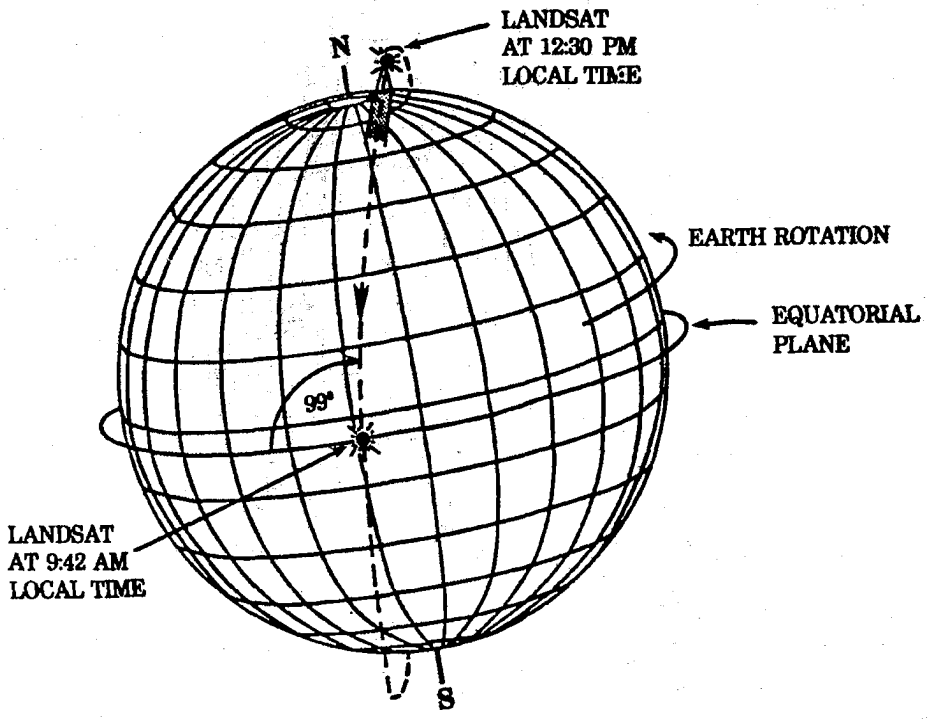


Figure 2–13 Inclination of the Landsat orbit to maintain sun synchronous orbit. (From Taranik, 1978.)

OFEK06711

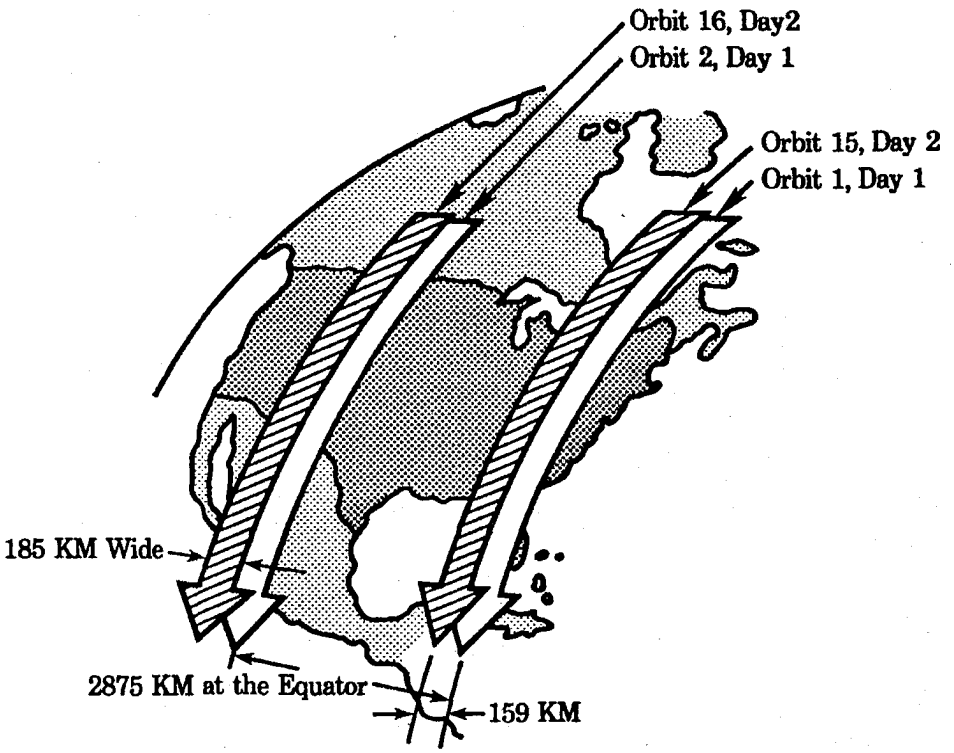


Figure 2–14 From one orbit to the next the position directly below the satellite moved 2875 km (1785 mi) at the equator as the earth rotated beneath it. The next day, 14 orbits later, it was approximately back to its original location, with orbit 15 displaced westward from orbit 1 by 159 km (99 mi) at the equator. (From Taranik, 1978.)

Figure 2–14 indicates how repeat coverage of a given geographic area was acquired. From one orbit to the next, a position directly below the spacecraft moved 2875 km (1785 mi) at the equator as the earth rotated beneath it. The next day, 14 orbits later, it was approximately back to its original location, with orbit 15 displaced westward from orbit 1 by 159 km (99 mi) at the equator. This continued for 18 days, after which orbit 252 fell directly over orbit 1 once again. Thus the Landsat sensor systems had the capability of observing the entire globe (except poleward of 81°) once every 18 days, or about 20 times a year. There was approximately 26 km (16 mi) of sidelap between successive orbits (Table 2–3). This overlap was a maximum at 81° north and south latitudes (about 85%) and a minimum at the equator (about 14%). This has proven useful for some stereoscopic analysis applications.

Sensor systems on board the Landsat satellites. In the context of this section on data acquisition, we are very interested in the type of sensors carried aloft by the Landsat satellites and the nature and quality of remote sensor data provided for earth resource investigations. The sensors included the multispectral scanner (MSS), return beam vidicon camera (RBV), and the thematic mapper (TM). However, the focus here will be only on the multispectral scanner and the thematic

OFEK06712

TABLE 2–3 Relationship Between Latitude of Landsat MSS and Image Sidelap

Latitude (deg)	Image sidelap (%)
0	14.0
10	15.4
20	19.1
30	25.6
40	34.1
50	44.8
60	57.0
70	70.6
80	85.0

mapper systems, as they continue to provide the more significant part of all remotely sensed imagery used for earth resource studies.

The *Landsat multispectral scanner system* was placed on each of the five Landsat satellites. It has collected more useful earth resource remote sensor data of the world than any other sensor. MSS data have been used extensively by scientists in a wide variety of applications. The MSS multidetector array and the scanning system are shown diagrammatically in Fig. 2–15. The scanning mirror oscillated through an angular displacement of ±2.89°. Because the angle of incidence was approximately equal to the angle of reflection, the scanned beam was approximately 11° wide. This resulted in a swath width of approximately 185 km (115 mi) for each orbit. Six parallel detectors sensitive to four spectral bands (channels) in the electromagnetic spectrum viewed the ground simultaneously: 0.5 to 0.6 μm (green), 0.6 to 0.7 μm (red), 0.7 to 0.8 μm (reflective infrared), and 0.8 to 1.1 μm (reflective infrared). These bands were designated as 4, 5, 6, and 7 because the return beam vidicon (RBV) subsystem included bands 1, 2, and 3. When not viewing the earth, the detectors were exposed to internal light and sun calibration sources. The spectral sensitivity of the bands is summarized in Table 2–4.

The instantaneous-field-of-view (IFOV) of each detector was square and resulted in a ground resolution element of approximately 79 × 79 m (67,143 ft²). The voltage analog signal from each detector was converted to digital value using an on-board analog-to-digital (A-to-D) converter. The data were quantized to 6 bits with a range of values from 0 to 63. These data were then rescaled to 7 bits (0 to 127) for three of the four bands in subsequent ground processing (i.e., bands 4, 5, and 6 were decompressed to a range of 0 to 127).

During each scan, the voltage produced by each detector was sampled every 9.95 μs. For one detector, approximately 3300 samples were taken along a 185-km line (Taranik, 1978). Thus the IFOV of 79 × 79 m became about 56 m on the ground between each sample (Fig. 2–16). The 56 × 79 m area is called a Landsat

OFEK06713

LANDSAT MULTISPECTRAL SCANNER SYSTEM

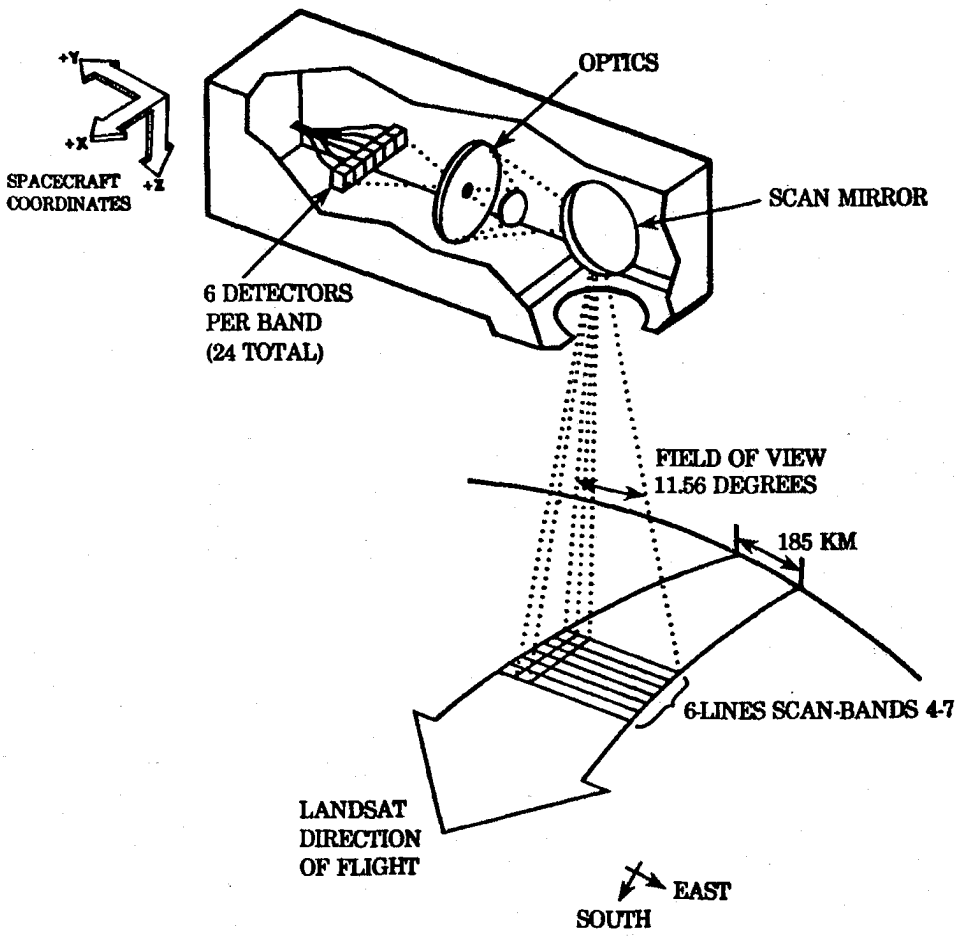


Figure 2–15 Major components of the Landsat multispectral scanner (MSS) system on Landsats 1, 2, 3, 4, and 5 (Landsat 3 also had a thermal infrared band). A bank of 24 detectors (six for each of the four bands) measures information from the earth from an instantaneous field of view (IFOV) of 79 × 79 m.

MSS picture element. Thus, although the measurement of landscape brightness was made from a 6241-m² area, each pixel was reformatted as if the measurement were made from a 4424-m² area (Fig. 2–16). Note the overlap of the areas from which brightness measurements were made for adjacent pixels.

The MSS scanned each line from west to east as the southward orbit of the spacecraft provided the along-track progression. Each MSS scene purchased from the EROS Data Center represented a 185 × 178 km parallelogram extracted from the continuous swath of an orbit and contained approximately 10% endlap. A typical scene contained approximately 2340 scan lines with about 3240 pixels per line or

OFEK06714

TABLE 2–4 Comparison of Multispectral Scanner and Thematic Mapper Sensor System Characteristics

Multispectral Scanner			Thematic Mapper		
Band number	Micrometers	Radiometric sensitivity (NEΔP)[a]	Band number	Micrometers	Radiometric sensitivity (NEΔP)
4[b]	0.5–0.6	0.57	1	0.45–0.52	0.8
5	0.6–0.7	0.57	2	0.52–0.60	0.5
6	0.7–0.8	0.65	3	0.63–0.69	0.5
7	0.8–1.1	0.70	4	0.76–0.90	0.5
8[c]	10.4–12.6	1.4K(NEΔT)	5	1.55–1.75	1.0
			7	2.08–2.35	2.4
			6	10.4–12.50	0.5 (NEΔT)
IFOV:		79 × 79 m for bands 4–7; 240 × 240 m for band 8		30 × 30 m for bands 1–5, 7; 120 × 120 m for band 6	
Data rate:		15 MB/s		85 MB/s	
Quantization levels:		6 bits, 64 levels		8 bits, 256 levels	
Earth coverage:		18 days, Landsats 1, 2, 3; 16 days, Landsats 4, 5		16 days	
Altitude:		919 km		705 km	

[a] The radiometric sensitivities are the noise-equivalent reflectance differences for the reflective channels expressed as percentages (NEΔP) and temperature differences for the thermal infrared bands (NEΔT).

[b] MSS bands 4, 5, 6, and 7 were renumbered bands 1, 2, 3, and 4 on Landsats 4 and 5.

[c] MSS band 8 was present only on Landsat 3.

about 7,581,600 pixels per channel. All four bands represented a data set of more than 30 million values.

Landsat MSS products have provided an unprecedented ability to observe large geographic areas within a single image. For example, it would require approximately 5000 conventional vertical aerial photographs obtained at a scale of 1:15,000 to equal the geographic coverage of a single Landsat MSS image (Table 2–5). This allows regional terrain analysis to be performed using one data source rather than a multitude of aerial photographs.

The first *thematic mapper (TM) sensor system* was launched on July 16, 1982, and the second on March 1, 1984. The TM is a scanning optical-mechanical sensor system that records reflected and emitted energy in the visible, reflective-infrared, middle-infrared, and thermal-infrared regions of the electromagnetic spectrum. It collects multispectral imagery that has higher spatial, spectral, and radiometric resolution than the Landsat MSS.

Detailed descriptions of the design and performance characteristics of the TM may be found in Engel and Weinstein (1983), Solomonson (1984), Holmes (1984) and Williams et al. (1984). The platform is shown in Fig. 2–17 and the sensor system configuration in Fig. 2–18. The telescope directs the incoming radiant flux obtained along a scan line through a scan line corrector (1) to the visible and near-infrared

OFEK06715

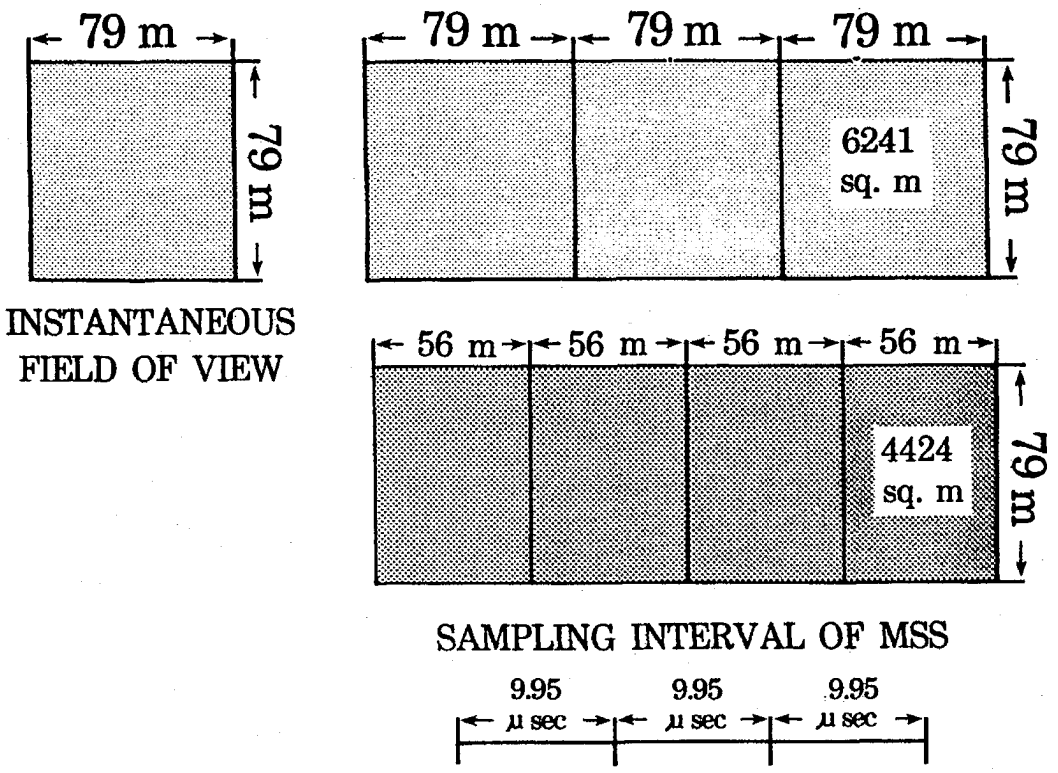


Figure 2–16 Relationship between the original 79 × 79 m IFOV of the Landsat MSS and the rate at which it was resampled (i.e., every 9.95 μs). This resulted in picture elements (pixels) that were 56 × 79 m in dimension on tapes purchased from the EROS Data Center at Sioux Falls, S.D.

TABLE 2–5 Comparison of Landsat MSS versus Aircraft Areal Coverage

Platform	Scale	Number of images or photos
Landsat MSS 1, 2, and 3 185 × 178 km = 32,930 km² 115 × 111 mi = 12,765 mi²	—	1
Low-altitude aircraft	1:15,000	5000
Low-altitude aircraft	1:30,000	1500
High-altitude aircraft	1:60,000	300
Commercial jet aircraft	1:90,000	150
Government civilian aircraft	1:120,000	85
Government military aircraft	1:250,000	30

OFEK06716

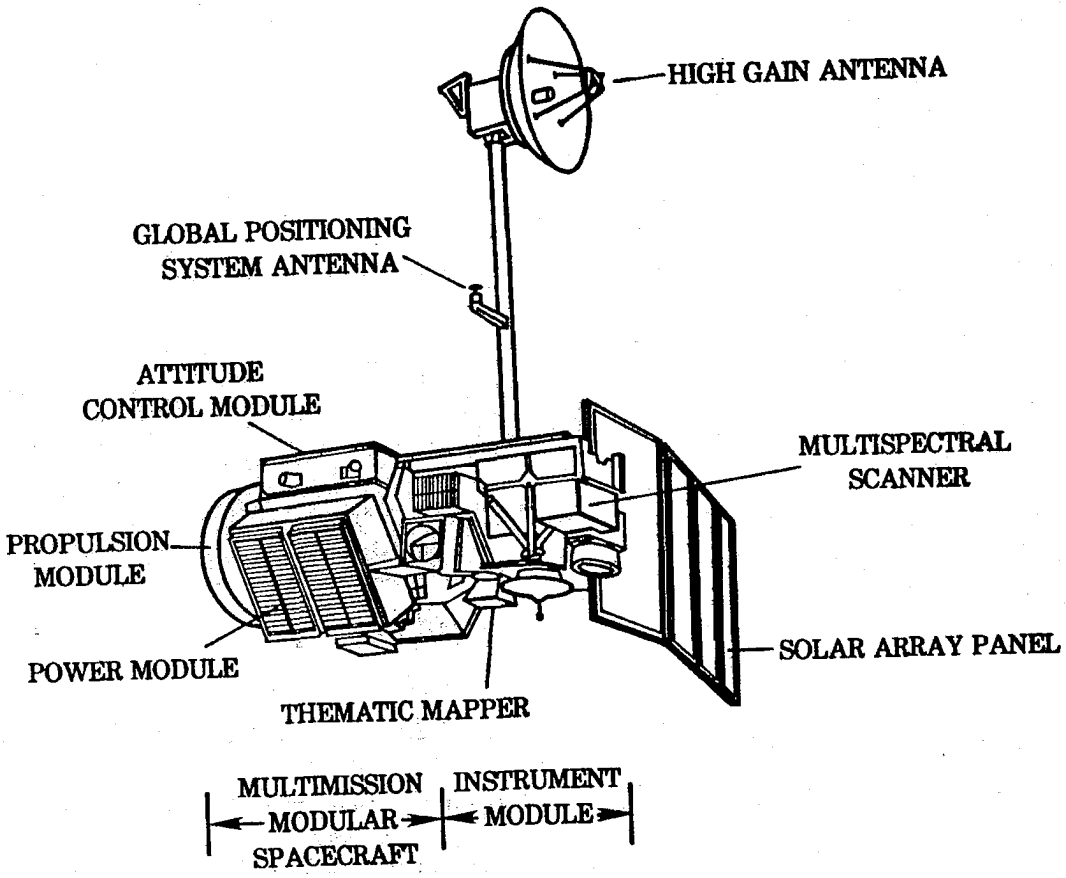


Figure 2–17 Landsat 4 and 5 platform with associated sensor and telecommunication systems.

focal plane, or (2) to the mid-infrared and thermal infrared cooled focal plane. The detectors for the visible and near-infrared bands (1 to 4) are four staggered linear arrays, each containing 16 silicon detectors. The two mid-infrared detectors are 16 indium–antimonide cells in a staggered linear array, and the thermal-infrared detector is a four-element array of mercury–cadmium–telluride cells.

The TM spectral bands represent important departures from the bands found on the traditional MSS also carried on board Landsats 4 and 5. The original MSS bandwidths were selected based on their utility for general vegetation inventories and general geologic studies. Conversely, most of the TM bands were chosen after years of analysis for their value in the discrimination of vegetation type and vigor, plant and soil moisture measurement, differentiation of clouds and snow, and identification of hydrothermal alteration in certain rock types (Table 2–6). For example, Fig. 2–19 demonstrates that the TM bands are situated to make maximum use of the dominant factors controlling leaf reflectance, such as pigmentation, leaf structure, and moisture content.

Landsat TM data have a ground projected instantaneous-field-of view (IFOV) of 30 × 30 m (Fig. 2–20). Classification accuracies become acceptable for most remote

OFEK06717

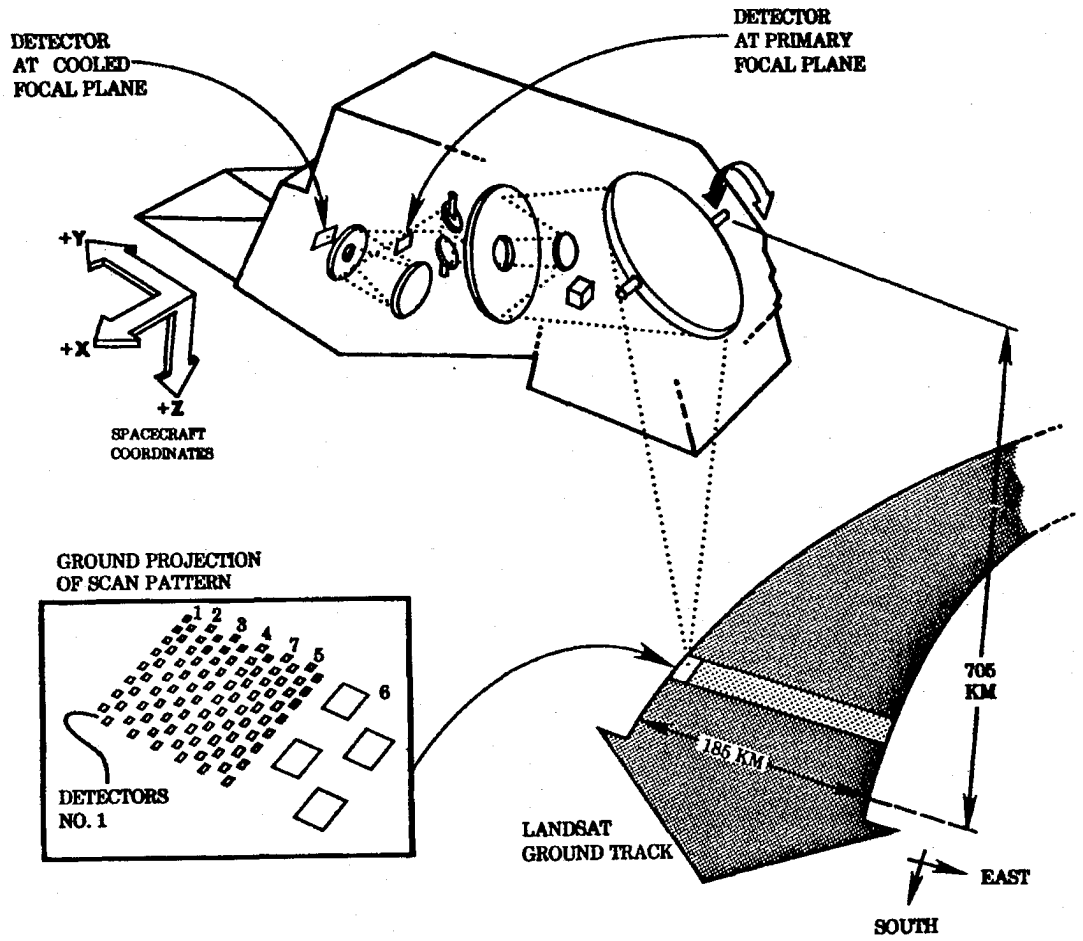


Figure 2–18 Landsat 4 and 5 thematic mapper sensor system. The sensor is sensitive to the seven bands of the electromagnetic spectrum summarized in Table 2–4. Six of the seven bands have a spatial resolution of 30 × 30 m while the thermal infrared band has a spatial resolution of 120 × 120 m. The telescope directs the incoming radiant energy obtained along a scan line through a scan line corrector to (1) the visible and near-infrared focal plane, or (2) to the mid-infrared and thermal-infrared cooled focal plane. The detectors for the visible and near-infrared bands (1 to 4) are four staggered linear arrays, each containing 16 silicon detectors. The two mid-infrared detectors are 16 indium antimonide cells in a staggered linear array, and the thermal-infrared detector is a four-element array of mercury–cadmium–telluride cells. (From Engel and Weinstein, 1983.)

sensing agriculture and forestry applications when field sizes are greater than 60 IFOVs in size. This corresponds to a square with about 8 IFOVs on a side. Thus effective classification accuracy should be obtained for fields 240 × 240 m in size if a 30 × 30 m ground IFOV is used. This should allow most of the field sizes in Canada, the United States, and the Soviet Union to be adequately sampled in the spatial domain as well as some information for fields in developing countries where field sizes are less than 240 × 240 m. The thermal-infrared band 6 has a spatial resolution of 120 × 120 m.

OFEK06718

TABLE 2-6 Characteristics of the TM Spectral Bands

Band 1: 0.45–0.52 μm (blue). Provides increased penetration of water bodies as well as supporting analyses of land use, soil, and vegetation characteristics. The shorter-wavelength cutoff is just below the peak transmittance of clear water, while the upper-wavelength cutoff is the limit of blue chlorophyll absorption for healthy green vegetation. Wavelengths below 0.45 μm are substantially influenced by atmospheric scattering and absorption.

Band 2: 0.52–0.60 μm (green). This band spans the region between the blue and red chlorophyll absorption bands and therefore corresponds to the green reflectance of healthy vegetation. Bands 1 and 2 combined provide the same spectral sensitivity as Kodak's Water Penetration Aerial Color Film (SO-224).

Band 3: 0.63–0.69 μm (red). This is the red chlorophyll absorption band of healthy green vegetation and represents one of the most important bands for vegetation discrimination. It is also useful for soil-boundary and geological boundary delineations. This band may exhibit more contrast than bands 1 and 2 because of the reduced effect of atmospheric attenuation. The 0.69-μm cutoff is significant because it represents the beginning of a spectral region from 0.68 to 0.75 μm, where vegetation reflectance crossovers take place that can reduce the accuracy of vegetation investigations.

Band 4: 0.76–0.90 μm (reflective-infrared). For reasons discussed above, the lower cutoff for this band was placed above 0.75 μm. This band is especially responsive to the amount of vegetation biomass present in a scene. It is useful for crop identification and emphasizes soil–crop and land–water contrasts.

Band 5: 1.55–1.75 μm (mid-infrared). This band is sensitive to the turgidity or amount of water in plants. Such information is useful in crop drought studies and in plant vigor investigations. In addition, this is one of the few bands that can be used to discriminate between clouds, snow, and ice, so important in hydrologic research.

Band 6: 2.08–2.35 μm (mid-infrared). This is an important band for the discrimination of geologic rock formations. It has been shown to be particularly effective in identifying zones of hydrothermal alteration in rocks.

Band 7: 10.4–12.5 μm (thermal infrared). This band measures the amount of infrared radiant flux emitted from surfaces. The apparent temperature is a function of the emissivities and true or kinetic temperature of the surface. It is useful for locating geothermal activity, thermal inertia mapping for geologic investigations, vegetation classification, vegetation stress analysis, and soil moisture studies.

The equatorial crossing time was changed from 9:30 A.M. for Landsats 1, 2, and 3 to 11:00 A.M. for Landsats 4 and 5. Higher sun angles reduce the amount of shadow in the imagery. This is useful for vegetation studies because shadows can modify spectral signatures particularly in forests and row crops. Conversely, low sun angles are useful for enhancing the geologic structure of the terrain. The 11:00 A.M. crossing time also, unfortunately, enhances the probability that cloud cover will be present over the study area and that increased amounts of particulate matter (e.g., smog) will be present in the atmosphere, especially over urban environments.

The transition from an approximately 919-km orbit to a 705-km orbit for Landsats 4 and 5 disrupted the continuity of Landsats 1, 2, and 3 MSS path and row designations in the Worldwide Reference System (WRS). Consequently, a separate WRS map is now required to select images obtained by Landsats 4 and 5. The lower orbit (approximately the same as the Space Shuttle) has also increased the amount

OFEK06719

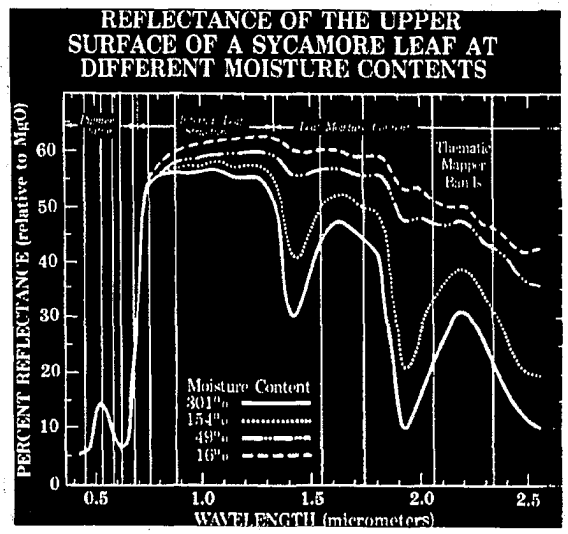


Figure 2–19 Progressive changes in percent reflectance for a sycamore leaf at varying oven-dry-weight moisture contents. The dominant factors controlling leaf reflectance and the location of six of the thematic mapper bands are superimposed.

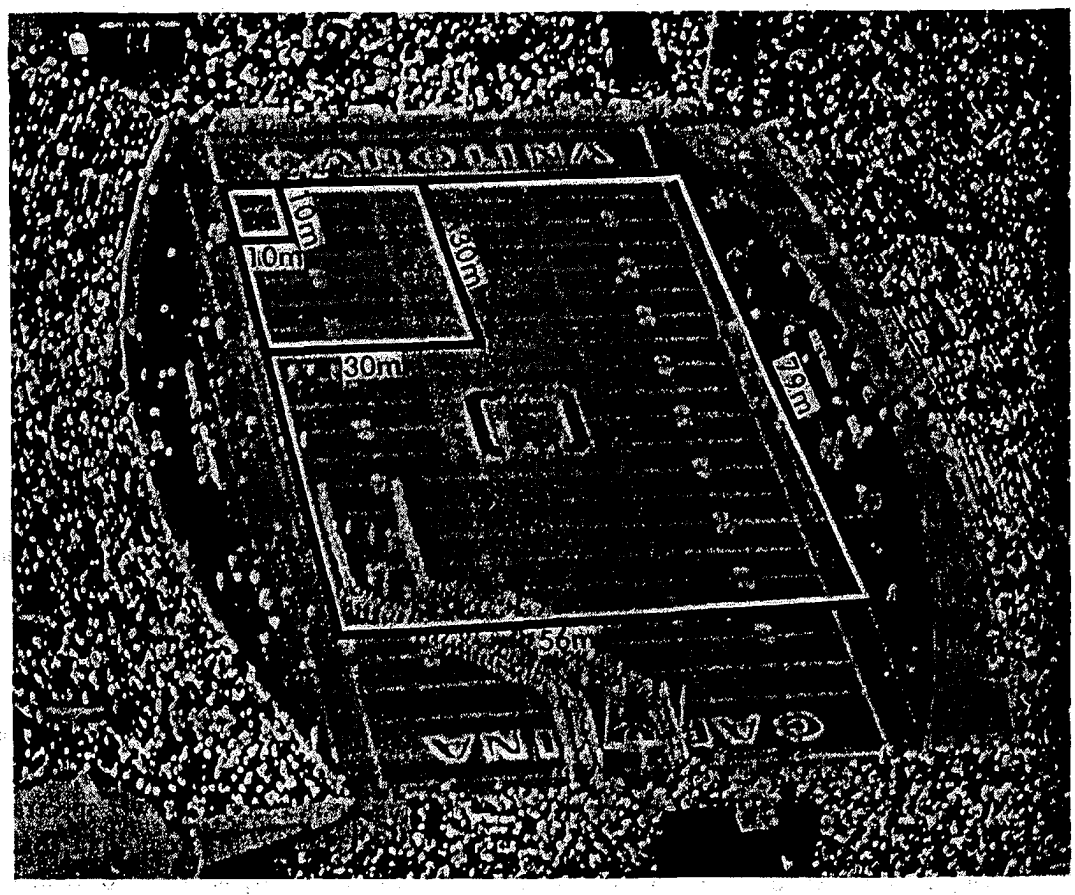


Figure 2–20 Diagram illustrating the spatial resolution of the Landsat MSS (56 × 79 m), the Landsat thematic mapper (30 × 30 m), and the SPOT panchromatic band (10 × 10 m).

OFEK06720

of relief displacement introduced into the imagery obtained over mountainous terrain. The new orbit also caused the period between repetitive coverage to change from 18 days to 16 days for both the MSS and TM data collected by Landsats 4 and 5.

There has been a substantial improvement in the level of quantization, from 6 bits to 8 bits per band (Table 2–4). This, in addition to a greater number of bands and a higher spatial resolution, has increased the data rate from 15 Mb/s to 85 Mb/s. This has required major modifications of ground receiving stations to handle the increased data flow.

Based on the improvements in spectral, spatial, and radiometric resolution, Solomonson (1984) suggests that

> from a wide variety of analyses and results, it appears that the TM can be described as being twice as effective in providing information as the Landsat MSS. This is based on its ability to provide twice as many separable classes over a given area as the MSS, numerically provide 2 more independent vectors or components in the data, or demonstrate through classical information theory that twice as much information exists in the TM data.

The SPOT Sensor System

The sensors on board the TIROS and NIMBUS satellites in the 1960s provided remotely sensed imagery with ground spatial resolutions of approximately 1000 × 1000 m and were the first to reveal the potential of space as a vantage point for earth resource observation. The multispectral Landsat MSS and TM sensor systems developed in the 1970s and early 1980s provided imagery with spatial resolutions of from 79 × 79 to 30 × 30 m. The SPOT sensor system developed by the French Centre National d'Études Spatiales (CNES) has a spatial resolution of 10 × 10 m and 20 × 20 m and provides several other innovations in remote sensor system design (Begni, 1982; CNES, 1984).

The SPOT satellite consists of two parts, the SPOT bus, which is a standard multipurpose platform, and a sensor system payload (Fig. 2–21). The payload of the first SPOT satellite consists of two identical high-resolution visible (HRV) sensor systems and a package comprising two magnetic-tape data recorders and a telemetry transmitter. The satellite operates in a circular sun-synchronous near-polar orbit (inclination of 98.7°) at an altitude of 832 km. The satellite passes overhead at the same solar time; the local clock time varies with latitude.

The HRV sensors can operate in two modes in the visible and reflective infrared portions of the spectrum: a *panchromatic* mode, corresponding to observation over a broad spectral band (similar to a typical black-and-white photograph), and a *multispectral* (color) mode, corresponding to observation in three narrower spectral bands (Table 2–7). Thus the spectral resolution is substantially different than that of the thematic mapper. The ground spatial resolution, however, is 10 × 10 m in the first case and 20 × 20 m in the second when the instruments are viewing directly below the spacecraft (i.e., at nadir). Radiant energy reflected from the terrain enters the

OFEK06721

TABLE 2–7 SPOT Sensor System Characteristics

Characteristics of the HRV sensors	Multispectral mode	Panchromatic mode
Spectral bands	0.50–0.59 μm 0.61–0.68 μm 0.79–0.89 μm	0.51–0.73 μm
Instrument IFOV	4.13°	4.13°
Ground sampling interval at nadir	20 × 20 m	10 × 10 m
Number of pixels per line	3000	6000
Ground swath width at nadir	60 km	60 km
Pixel quantization	8 bits	6 bits DPCM[a]
Image data bit rate	25 Mb/s	25 Mb/s

[a] DPCM (digital pulse code modulation) is a mode of data compression that does not degrade the radiometric accuracy of the 8-bit or 256-gray-level image data.

HRV via a plane mirror and is then projected onto two charge-coupled-detector (CCD) arrays (Fig. 2–21). Each CCD array consists of 6000 detectors arranged linearly (refer to Fig. 2–8). This is commonly referred to as a "push-broom" scanner since it images a complete line of the ground scene in the cross-track direction in one "look" as the sensor system progresses down-track. This capability breaks tradition with the Landsat MSS and TM sensors in that no mechanical scanning takes place.

When looking directly at the terrain beneath the sensor system, the two HRV instruments may be pointed to cover adjacent fields each 60 km (Fig. 2–21). In this configuration the total swath width is 117 km and the two fields overlap by 3 km. Since the distance between adjacent ground tracks at the equator is approximately 108 km, complete earth coverage can be obtained with this fixed setting of instrument fields of view.

However, it is also possible to selectively point the mirror to off-nadir viewing angles through commands from the ground station. In this configuration it is possible to observe any region of interest within a 950-km-wide strip centered on the satellite ground track (i.e., the observed region need not necessarily be centered on the ground track) (Fig. 2–22). The width of the swath actually observed varies between 60 km for nadir viewing and 80 km for extreme off-nadir viewing.

If the HRV instruments were only capable of nadir viewing the revisit frequency for any given region of the world would be 26 days. This interval is often unacceptable for the observation of phenomena evolving on time scales ranging from several days to a few weeks, especially where the cloud cover hinders the acquisition of usable data. During the 26-day period separating two successive SPOT satellite passes over a given point on the earth, and taking into account the steering capability of the instruments, the point in question could be observed on seven different passes if it were on the equator and on 11 occasions if at a latitude of 45° (Fig. 2–23). A given

0924

OFEK06722

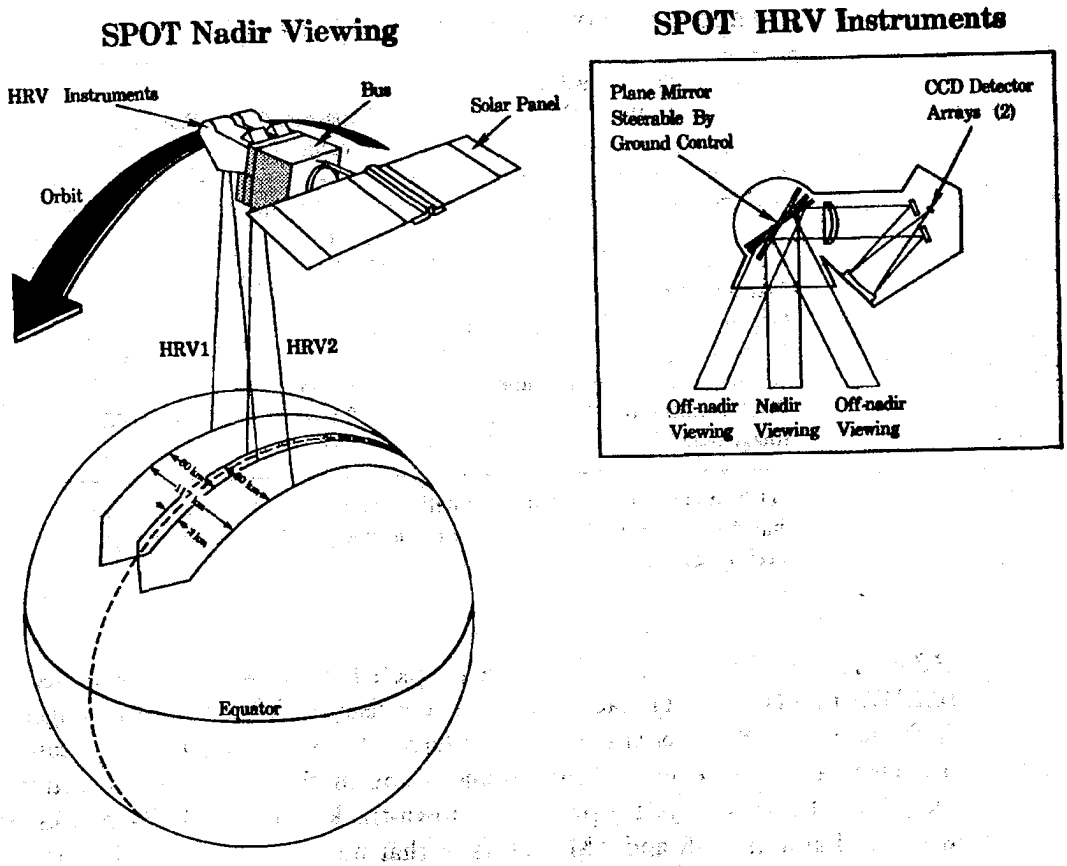


Figure 2–21 The SPOT satellite consists of two parts, the SPOT bus, which is a standard multipurpose platform, and the sensor system payload. The payload of the first SPOT satellite consists of two identical high-resolution visible (HRV) sensor systems. Radiant energy from the terrain enters the HRV via a plane mirror and is then projected onto two charge-coupled-detector (CCD) arrays. Each CCD array consists of 6000 detectors arranged linearly. This results in a spatial resolution of 10 × 10 or 20 × 20 m, depending on the mode in which the sensor is being used (refer to Table 2–7). (Courtesy of SPOT Image Corporation.)

region may be revisited on dates separated alternatively by 1 and 4 (or occasionally 5) days.

The SPOT sensors may also acquire stereoscopic pairs of images for a given geographic area (Fig. 2–24). Two observations may be made on successive days such that the two images are acquired at angles on either side of the vertical (CNES, 1984). In such cases, the ratio between the observation base (distance between the two satellite positions) and the height (satellite altitude) is approximately 0.75 at the equator and 0.50 at a latitude of 45°. The stereoscopic imagery obtained using this base–height ratio is of use in photogrammetric applications such as cartographic work at 1:100,000 scale and map updating at 1:50,000.

OFEK06723

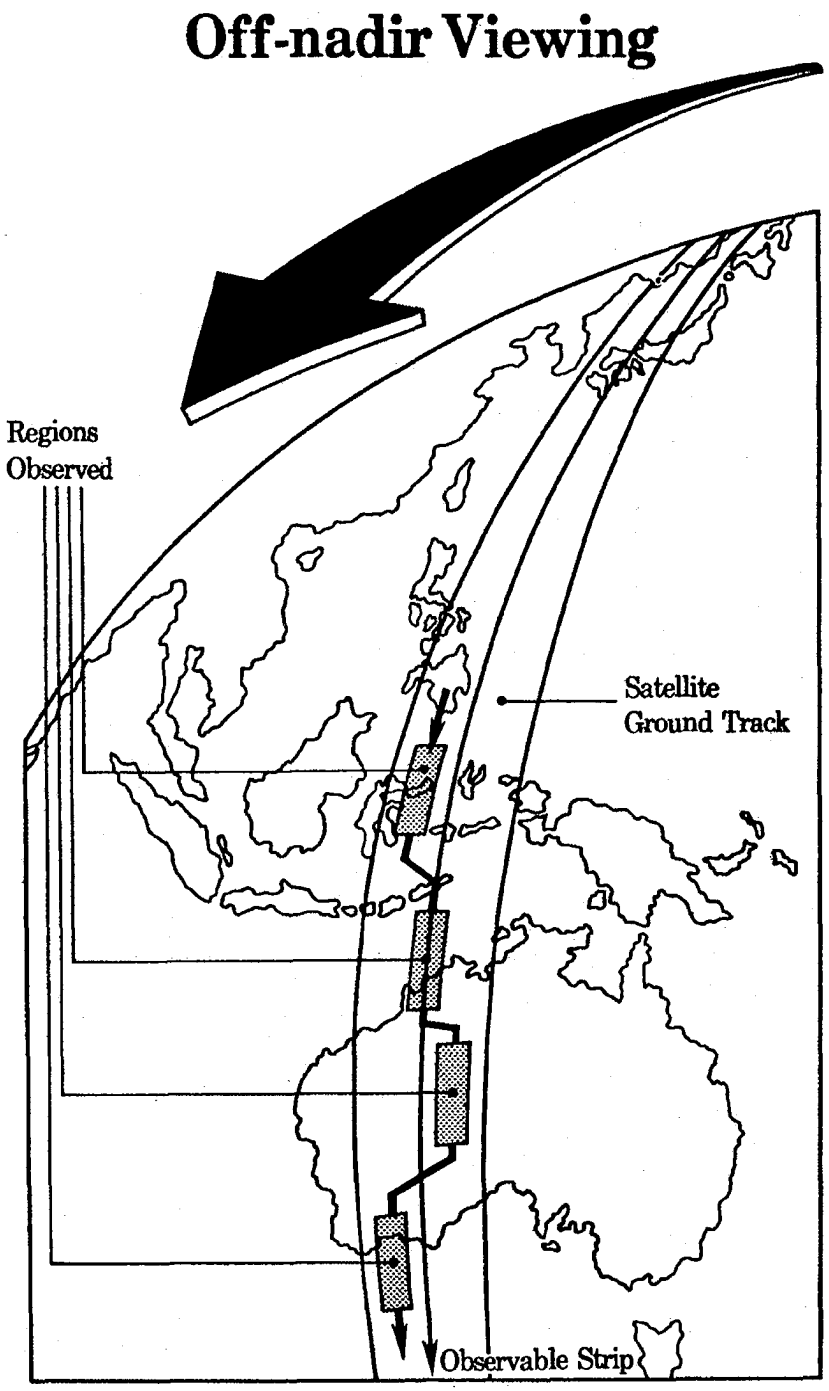


Figure 2–22 If the HRV instruments were only capable of viewing the terrain directly below the spacecraft (i.e., at nadir), the revisit frequency for any given region of the world would be 26 days. However, SPOT is pointable. This means that it can be used to view areas that are not directly below the aircraft (i.e., off nadir). This is very useful for collecting information in the event of a natural or human-made disaster when the satellite track is not optimum, or for collecting stereoscopic imagery. (Courtesy of SPOT Image Corporation.)

OFEK06724

SPOT Off-nadir Revisit Capabilities

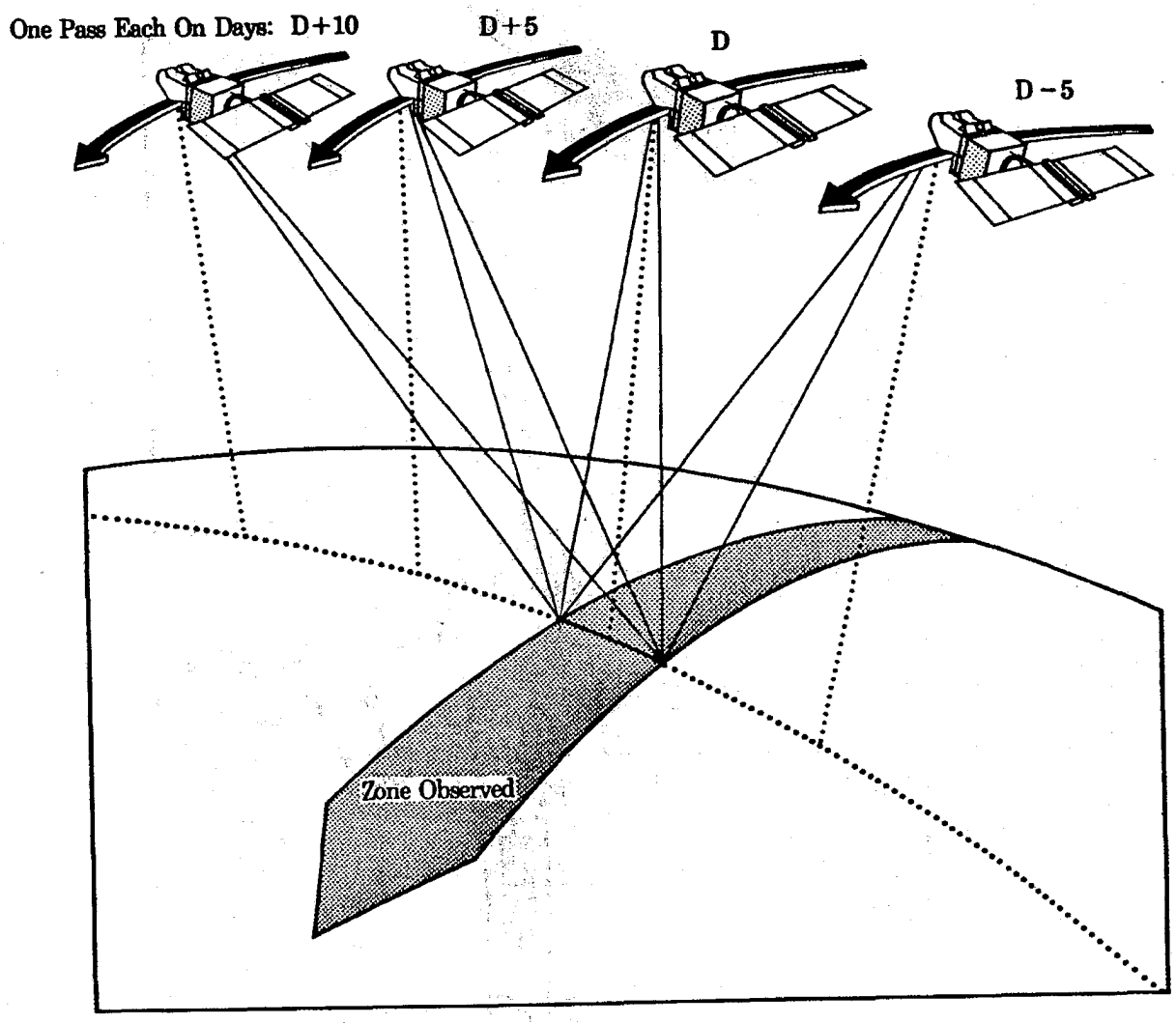


Figure 2–23 During the 26-day period separating two successive SPOT satellite overpasses, a point on the earth could be observed on seven different passes if it is at the equator and on 11 occasions if at a latitude of 45°. A given region can be revisited on dates separated alternatively by 1, 4, and occasionally 5 days. (Courtesy of SPOT Image Corporation.)

The SPOT satellite was launched in early 1986. Prior to launch, SPOT simulation imagery were acquired using a modified MSS carried on board suborbital aircraft. A SPOT simulation image of the Washington, D.C., area is shown in Fig. 2–25 (color section). The high spatial resolution in this image provides enough detail that analysts may visually interpret the data to extract useful urban thematic information. In addition, however, the data are provided in digital format for image processing purposes.

OFEK06725

Stereoscopic Viewing Capabilities

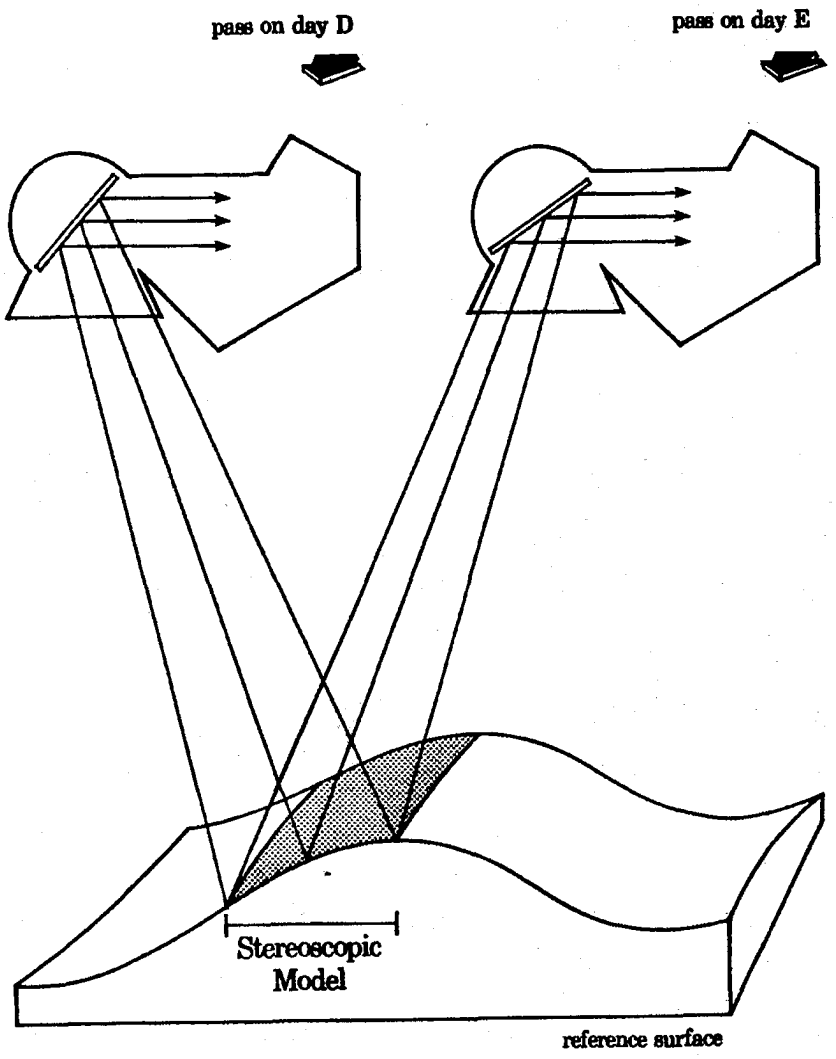


Figure 2-24 Two observations can be made on successive days such that the two images are acquired at angles on either side of the vertical, resulting in stereoscopic imagery. Such imagery can be used to produce topographic and planimetric maps. (Courtesy of SPOT Image Corporation.)

Aircraft Multispectral Scanners

Orbital sensors such as the Landsat MSS and TM collect data on a repetitive cycle and at set spatial and spectral resolutions. Often it is necessary to acquire remotely sensed data at specific times that do not coincide with the scheduled overpasses and at perhaps different spatial and spectral resolutions. For example, it may not be practical to wait for the next Landsat or SPOT overpass (even though pointable)

OFEK06726

when an oil spill or severe flood has occurred. Rapid collection and analysis of high-resolution remotely sensed data may be required to allocate resources and personnel. When such conditions occur or when a sensor configuration different from the Landsat or SPOT sensors is needed, agencies and companies often use a multispectral scanner placed on board an aircraft to acquire remotely sensed data. For example, Kahle and Goetz (1983) describe a new six-channel aircraft MSS developed to extract mineral signature information at wavelengths between 8 and 12 μm (refer to Table 2–2). Preliminary results show that igneous rock units may be identified from their free silica content, and that carbonate as well as clay-bearing units are readily separable on the digitally processed images.

The sensor system characteristics of a typical airborne MSS are shown in Fig. 2–26. Its principles of operation and components are displayed in Fig. 2–27. Radiant flux reflected or emitted from the terrain is collected by the scanning optical system and projected onto a *dichroic grating*. The grating separates the reflected radiant flux from the emitted radiant flux. Energy in the reflective part of the spectrum, including ultraviolet, blue, green, red, and reflective infrared, is directed from the grating to a prism (or refraction grating) that further separates the energy into specific bands. For example, the sensor system shown in Fig. 2–27 is sensitive to one ultraviolet band, two blue, two green, two red, and three reflective infrared bands. At the same time, the emitted incident energy is separated from the reflective incident energy.

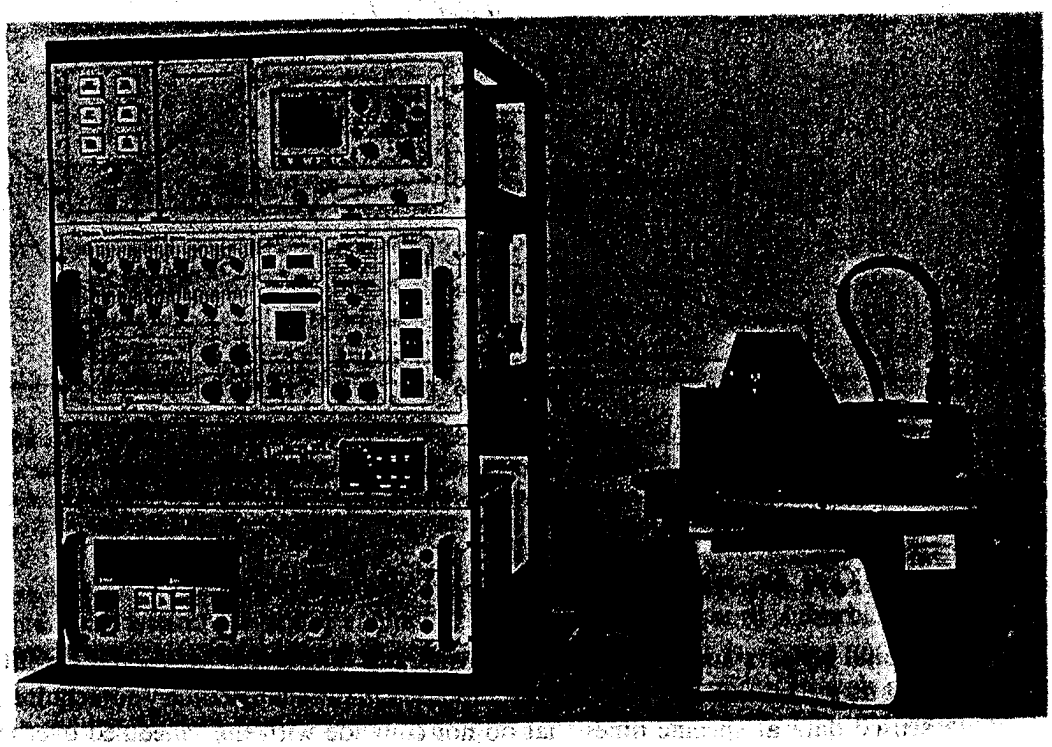


Figure 2–26 AADS 1260 multispectral scanner (MSS) and associated electronics that are carried on-board the aircraft during data collection. (Courtesy of Daedalus Enterprises, Inc.)

OFEK06727

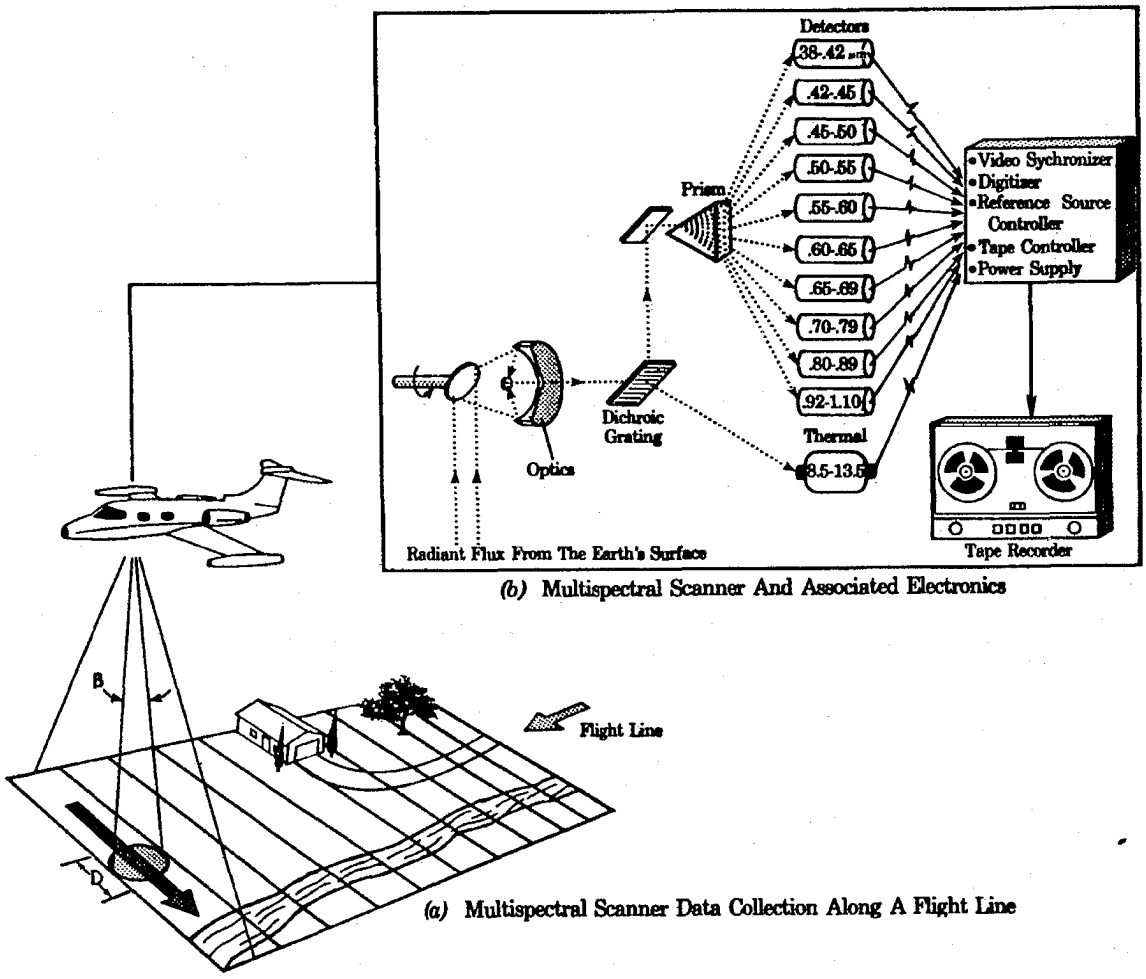


Figure 2–27 Aircraft multispectral scanner (MSS) sensor system components and characteristics of data collection along a flight line. The diameter of the circular ground area viewed by the sensor, D, is a function of the instantaneous field of view, β, of the scanner and the altitude above ground level of the aircraft, H, at the time of data collection. Radiant flux from the earth's surface is passed from the optical system onto a dichroic grate which sends the various wavelengths of light to detectors which are continuously monitored by the sensor system electronics.

The independent bands of energy are then focused onto a bank of detectors situated behind the grating and the prism. In this example, there are 10 detectors for the reflective bands and a single detector for the emitted band. The detector that records the emitted energy is usually cooled by a Dewar of liquid nitrogen or some other substance.

The signals recorded by the detectors are amplified by the system electronics and recorded on a multichannel tape recorder. If the emphasis is on visual analysis of the MSS data, it may be recorded on an analog recorder and converted directly into hard-copy imagery during or after the flight. If the data are to be digitally processed, it is necessary to convert the analog scanner output into a numerical

OFEK06728

(digital) format. This A-to-D conversion requires that the data collected in the several bands are precisely synchronized. Such data are recorded on high-density digital tape (HDDT) on board the aircraft. Later, the HDDT data are converted into a computer-compatible-tape (CCT) format suitable for digital image processing.

The flight altitudes for aircraft MSS surveys are determined by evaluating the size of the desired ground resolution element (or pixel) and the size of the study area. Basically, the diameter of the circular ground area viewed by the sensor, D, is a function of the instantaneous field of view, β, of the scanner and the altitude above ground level, H, where

$$D = H \times \beta \tag{2-1}$$

If, for example, the IFOV for the scanner is 2.5 mrad, the ground size (in meters or feet) of the pixels is a product of the IFOV (0.0025) and the altitude AGL (in meters or feet). Table 2–8 presents selected flight altitudes and corresponding pixel sizes at nadir for an IFOV of 2.5 mrad.

TABLE 2–8 Aircraft MSS Flight Altitudes, Pixel Sizes, and Flight Line Coverage

Flight altitude AGL (m)	Pixel size[a] (m)	Ground coverage per flight line[b] (m)
1,000	2.5	1,860
2,000	5.0	3,720
4,000	10.0	14,880
16,000	40.0	29,760
50,000	125.0	93,000

[a] Assumes an IFOV of 2.5 mrad.

[b] Swath width equals 1.86 × flight altitude.

The following factors should be considered when collecting aircraft MSS data (Hawley, 1982):

1. When looking for a small phenomenon, the pixel size must be smaller than the objects being resolved. The chance of resolving an object similar in size to the pixel is marginal.
2. The size of the survey area is a factor when considering survey altitude and pixel size. At lower altitudes, the better resolution may be outweighed by the fact that more flight lines are required to cover the area compared to more efficient coverage at higher altitudes with larger pixels. The pixel size and the size of the survey are considered, objectives are weighed, and a compromise is reached.

OFEK06729

The field of view of the MSS optical system and the altitude AGL dictate the width of a single flight line of coverage. In this example, the swath width is determined by multiplying 1.86 times the altitude AGL (Table 2–8).

To collect aircraft MSS data, it is necessary that there be a relationship between the MSS scan speed (scans/s) and the on-board tape recorder speed (in./s). This is computed based on the use of a *v/h ratio*. The *v/h* ratio is calculated by dividing the flight altitude AGL into the velocity of the aircraft. The *v/h* ratios for the Daedalus DS-1260 scanner are shown in Table 2–9. The range of *v/h* values in the table is less than the maximum *v/h* range that allows contiguous scanner coverage. That is, at maximum *v/h* ranges, the scan lines are adjacent with *no* scan line overlap. It is best to work well within the *v/h* ranges and below the *v/h* maximum to guarantee some overscanning. This ensures adequate ground coverage. The amount of overscan is determined by dividing the maximum *v/h* value by the actual *v/h* value. Unity means that there is no overscan. A value of less than 1 indicates underscanning and that adequate coverage is not being obtained. A value greater than 1 means that there is some degree of overscanning. The exact amount of overscan is determined by subtracting 1 from the quotient. Thus there is a fine balance between too much overscanning taking place, which provides redundant information for a given scan line on the ground, and underscanning, which may result in gaps between scan lines.

TABLE 2–9 Daedalus DS-1260 MSS *v/h* Ratios and Recommended Scan and Tape Speed

Recommended *v/h* range	Max. *v/h*	Scan speed (scans/s)	Tape speed (in./s)
0.00–0.02	0.031	12.5	7.5
0.02–0.05	0.062	25.0	15.0
0.05–0.12	0.125	50.0	30.0
0.12–0.25	0.250	100.0	60.0

DIGITAL IMAGE DATA FORMATS

Having described how the Landsat MSS, TM, SPOT, and airborne MSS remote sensors acquire the data, it is important to identify some of the most prevalent formats for storing the digital data. These include: (1) band sequential (BSQ), (2) band interleaved by line (BIL), (3) band interleaved by pixel (BIP), and (4) run-length encoding. The formats will be reviewed in the context of a typical Landsat MSS tape. It should be noted, however, that each of these formats is usually preceded on the digital tape by "header" and/or "trailer" information, which consists of ancillary data about the date, altitude of the sensor, attitude, sun angle, and so on. Such information is useful when geometrically or radiometrically correcting the data.

OFEK06730

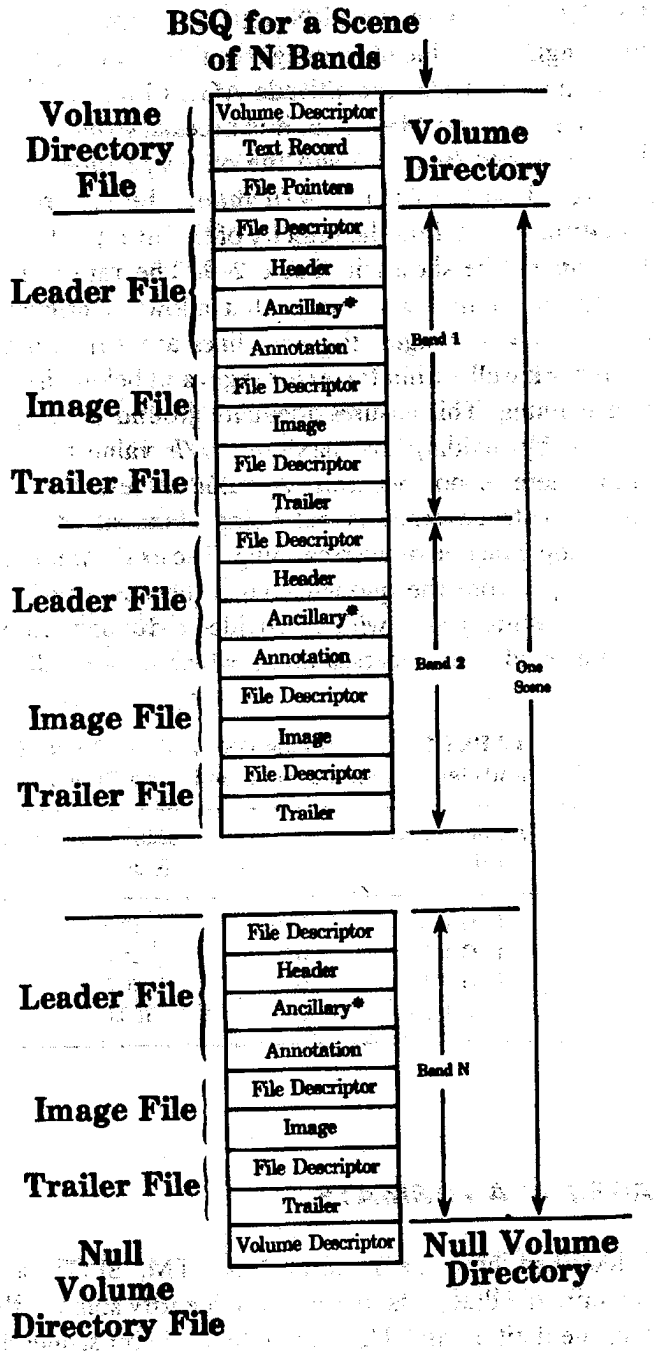


Figure 2–28 Format for a band sequential (BSQ) Landsat tape.

OFEK06731

Digital Landsat MSS data obtained from the EOSAT Company, Sioux Falls, South Dakota, as a computer-compatible tape (CCT) is in a format that has been converted from the original wide-band videotape data recorded by the various world-wide receiving stations. Each Landsat MSS image is composed of a sequence of scan lines and a variable number of pixels. The data are normally recorded on nine-track CCTs with data density on the tape of 800, 1600, or 6250 bits per inch (bpi).

Band Sequential Format

The band sequential format requires that all data for a single band covering the entire scene be written as one file. Thus if one wanted the area in the center of a scene in four bands, it would be necessary to read into this location in four separate files to extract the desired information. Many researchers like this format because it is not necessary to read "serially" past unwanted information if certain bands are of no value. The format for typical BSQ tape is summarized in Fig. 2–28. The number of tapes may be dependent on the number of bands provided and on whether the imagery has been geometrically corrected at the EOSAT Company. Program LAND-READ in Appendix A will extract a user-specified subscene from a band sequential Landsat MSS scene. Program TMREAD will perform the same function on thematic mapper imagery obtained in BSQ format from the EOSAT Company.

Band Interleaved by Line Format

In this format, the data for the bands are written line by line onto the same tape (i.e., line 1 band 1, line 1 band 2, line 1 band 3, line 1 band 4, etc.). It is a useful format if all the bands are to be used in the analysis (Fig. 2–29). If some bands are not of interest, the format is inefficient since it is necessary to read serially past unwanted data.

Band Interleaved by Pixel Format

Every element in the matrix has four pixel values (one from each spectral band) placed one after the other [i.e., pixel (1,1) band 1, pixel (1,1) band 2, pixel (1,1) band 3, pixel (1,1) band 4, etc.]. Again, this is a practical data format if all bands are to be used.

Run-Length Encoding Format

This is a data structure borrowed from cartographic theory which conserves storage space (Monmonier, 1982). It is a band sequential format that keeps track of both the brightness value and the number of times the brightness value occurs along a given scan line. For example, if a body of water were encountered with brightness values of 10 for 60 pixels along a scan line, this might be stored in the computer in integer (213) format as 060010, meaning that the following 60 pixels will each have

OFEK06732

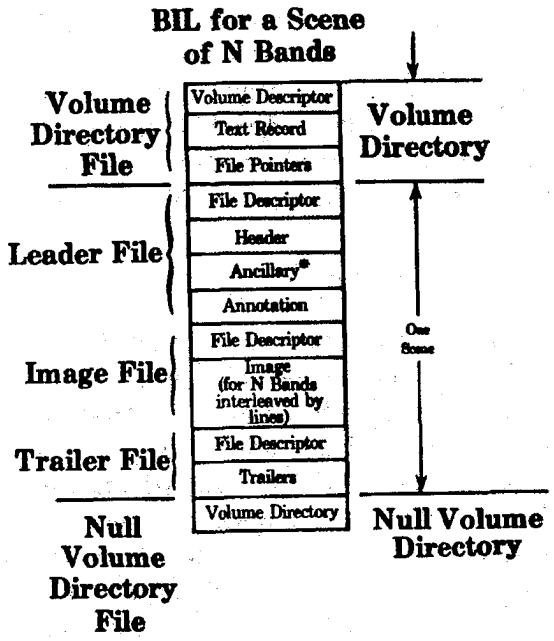


Figure 2–29 Format for a band interleaved by line (BIL) Landsat tape.

a brightness value of 10. Obviously, storing the two values 60 and 10 would require far less memory on disk or tape than storing 60 number 10s. However, if the data are exceptionally heterogeneous, with very few similar brightness values, this format is no better than the others. It is not possible at this time to purchase Landsat products in run-length-encoded format from the EOSAT Company. Rather, analysts often reformat the original data on the CCT into a run-length-encoded format for specific applications.

REFERENCES

BEGNI, G. 1982. "Selection of the Optimum Spectral Bands for the SPOT Satellite," *Photogrammetric Engineering and Remote Sensing*, Vol. 48, pp. 1613–1620.

CNES. 1984. *SPOT Satellite-Based Remote Sensing System*. Toulouse, France: Centre National d'Études Spatiales.

ENGEL, J. L., and O. WEINSTEIN. 1983. "The Thematic Mapper—An Overview." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. GE-21, pp. 258–265.

HAWLEY, D. L. 1982. *A Field Guide for Scanner and Photographic Missions*. Publication 1183–1818 (September). Las Vegas, Nev.: E.G.G. Energy Measurements Group, pp. 8–10.

OFEK06733

HOLMES, R. A. 1984. "Advanced Sensor Systems: Thematic Mapper and Beyond," *Remote Sensing of Environment*, Vol. 15, pp. 213–221.

HOLMQUIST, K. W., D. S. ANDREWS and A. W. VOSS. 1984. "Digitizing National High Altitude Photography," *Technical Papers, 50th Annual Meeting of the American Society of Photogrammetry*, pp. 694–699.

KAHLE, A. B., and A. GOETZ, 1983. "Mineralogic Information from a New Airborne Thermal Infrared Multispectral Scanner," *Science*, Vol. 222, pp. 24–27.

LANDGREBE, D. A. 1983. "Land Observation Sensors in Perspective," *Remote Sensing of Environment*, Vol. 13, pp. 391–402.

MONMONIER, M. S. 1982. *Computer-Assisted Cartography: Principles and Prospects*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., pp. 36–40.

NOAA. 1975 to 1985. *Landsat Data Users Notes*. NOAA Landsat Customer Services, Mundt Federal Building, Sioux Falls, SD 57198.

SOLOMONSON, V. V. 1984. "Landsat 4 and 5 Status and Results from Thematic Mapper Data Analyses," *Proceedings*, Machine Processing of Remotely Sensed Data Symposium. West Lafayette, Ind.: Laboratory for the Applications of Remote Sensing, pp. 13–18.

SOUTHWORTH, C. S. 1985. "Characteristics and Availability of Data from Earth-Imaging Satellites," *U.S. Geological Survey Bulletin #1631.* Washington, D.C. U.S. Geological Survey, 102 pages.

TARANIK, J. V. 1978. "Characteristics of the Landsat Multispectral Data System," U.S. Dept. of the Interior *Open File Report 78-187*, Sioux Falls, S.D.

U.S. GEOLOGICAL SURVEY. 1979. *Specifications for Inter-agency High Altitude Aerial Photography* (79–30), Reston, Va.: U.S. Geological Survey.

WILLIAMS, D. L., J. R. IRONS, B. L. MARKHAM, R. F. NELSON, D. L. TOLL, R. S. LATTY, and M. L. STAUFFER. 1984. "A Statistical Evaluation of the Advantages of Landsat Thematic Mapper Data in Comparison to Multispectral Scanner Data," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. GE-22, May issue.

0936

OFEK06734

3

Image Processing System Considerations

INTRODUCTION

Once the remotely sensed data are in a digital format, it is possible to analyze them using a digital image processing system which can hopefully extract meaningful information from the imagery. This requires that the analyst understands the fundamental characteristics of digital image processing systems.

Three approaches are often used to perform digital image processing in educational and research environments (Jensen, 1983). The first is the mainframe approach, where all analysts utilize a mainframe computer ($\geq$ 32-bit) in batch or interactive mode. They usually work with alphanumeric overprint or line plotter output and only occasionally have the opportunity to view and interactively analyze the remote sensor data on a high-resolution black-and-white or color monitor. Also, the typical afternoon slowdown caused by too many users on the mainframe can become unbearable (Williams et al., 1983; Eyton, 1983).

The second approach involves the use of a minicomputer (16- to 32-bit), an image processor, and a high-resolution color monitor which provide an ideal image processing environment (Ince, 1982). Unfortunately, such systems are usually expensive and often only one person can interact with the system and high-resolution monitor at one time.

Finally, there exist microcomputer-based (16- to 32-bit) image processing systems. Numerous relatively inexpensive microcomputer-based image processing systems may now be acquired for use as multiple workstations in a digital image processing laboratory environment.

Characteristics of some of the most widely used commercial and publicly available image processing systems are summarized in Tables 3–1 and 3–2, including

0937

OFEK06735

TABLE 3–1 Some Commercially Available Digital Image Processing Hardware and Software for Earth Resource Mapping, Including Several Geographic Information Systems

System	Typical CPU and image processor	Pixels (rows × cols.)	Bits/ pixel	Displayable colors vs. palette	Compiler(s)	Image processing functions[a]	Cost[b]
AIPE	Apple II, 48K	40 × 40	4	16/16	BASIC	E,I,J	B.R.
APPLEPIPS	Apple II, 48K	40 × 40	4	16/16	BASIC	E,I,H,P,Z	0.5K
		96 × 140	3	6/6	Assembly		
GRIDAPPLE	Apple II, 48K	32 × 32	4	16/16	Pascal	E,T,Z	5.0K
MICROPIPS	≥IBM PC, 128K	200 × 320	4	4/16	BASIC	A,E-K,P,Z	0.5K
		160 × 100		16/16			
MYERS	≥IBM PC, 128K	200 × 320	4	4/16	BASIC A Assembly	E,J,P	B.R.
SPECDAT	≥IBM PC, 128K VIP, 96K	240 × 256	12	4K/16M	C BASIC	A,C-K,Z	6–10K
ERDAS-PC	IBM AT, 512K #9, 917K	512 × 512	32	16M/16M	FORTRAN Assembly	Y,Z	12–20K
DECISION-IMAGES	IBM AT 512K #9, 917K	512 × 512	24	16M/16M	Pascal	Y,Z	12–20K
TERRA-MAR	IBM AT, 512K N.A.	512 × 512	24	16M/16M	FORTRAN	Y,Z	20–42K
ERDAS-2400	DEC VAX, 512K GOULD, 917K	512 × 512	26	16M/16M	FORTRAN Assembly	Y,Z	75–100K 50–125K
COMTAL	≥DEC 11/24, 512K COMTAL, 917K	512 × 512	28	16M/16M	FORTRAN Assembly	Y with VICAR/ELAS	
IDIMS	H-P 3000, 256K DeAnza 1M	512 × 512	32	16M/16M	FORTRAN Assembly	Y	100–400K
I²S	≥DEC 11/24, 256K Model 75, 917K	512 × 512	28	16M/16M	FORTRAN Assembly	Y with System 575	50–100K
LogE-ISI	DPS11/24, 128K VDI 100, 512K	512 × 512	16	64K/64K	FORTRAN Assembly	W,Z	80–250K
ESRI-ARC-INFO-GRID-GRIDTOPO	PRIME 50–9950	—	—	—	FORTRAN Assembly	T,U	95K
ODYSSEY	VAX 11/780	—	—	—	FORTRAN	T,U	10K

[a] See Table 3–3 for function codes.
[b] B.R., by request.
Source: Revised Jensen (1984).

0938
OFEK06736

TABLE 3–2 Some Publicly Available Digital Image Processing Software for Earth Resource Mapping, Including Several Geographic Information Systems[a]

System	Typical CPU image processor	Language and mode[b]	Radiometric and geometric correction	Enhancement	Display and analysis	GIS	Cost[c]
MINI-LARSYS	Apple II, 48K	BASIC (I)	—	E	J,P	—	B.R.
IMAGES	IBM 370	BASIC (I)	A,B	E-N	P	—	0.2K
RIPS	Cromemco, 64K SPECDAT, 96K	FORTRAN (I) Assembly	A	C-K	P,Z	—	B.R.
HOTLIPS	CalComp, 64K CAT, 132K	FORTRAN (I) Assembly	A-C	C-O	P-S	—	B.R.
ELAS	Perkin-Elmer or Data General	FORTRAN (I) Assembly	A,B	C-M	P-S	T	3.7K
KARS	Honeywell 66	FORTRAN (I)	A,B	E-L	P,Q	T	B.R.
EYTON	IBM 3081 and SAS	FORTRAN (I)	A,B	E-L	P,Q	—	B.R.
ORSER-OCCULT	IBM 370	FORTRAN (I) Assembly	A,B	C-O	P-S	T	2.2K
VICAR-IBIS	IBM 370	FORTRAN (I) Assembly	A,B	C-O	P-S	T	2.4K
LARSYS-TAMU	IBM 370	FORTRAN (I) Assembly	A,B	C-M	P-S	T	2.6K
GYPSY	VAX 11/780 or IBM 370	FORTRAN (I) Assembly	A,B	C-O	P-S	—	B.R.
AOIPS-CLASSPAK	DEC 11/70	FORTRAN (I) Assembly	A,B	C-O	P,Q	—	2.4K
ERIPS	IBM 370	FORTRAN (I) Assembly, PL/1	A,B	C-M	P,Q	—	B.R.
SMIPS	IBM 370	FORTRAN (B)	A,B	C-M	P,Q	—	5.0K

[a] See Appendix B for the address of people, companies, or government agencies that supply these programs.

[b] I, Interactive image processing; B, batch image processing.

[c] B.R., by request.

Source: Revised Jensen (1984).

0939

OEEK06737

information on their hardware/software configurations and cost where applicable (Welch et al., 1983; Jensen, 1984). Addresses of public and commercial vendors of digital image processing hardware and software are given in Appendix B.

The image processing functions typically performed by such systems are summarized in Table 3–3 (Ince, 1982). There is a positive relationship between the functions performed, the speed at which the image processing is performed, and the cost. Most of the minicomputer-based image processing systems cost from $20,000 to $150,000 and provide all the essential digital image processing functions (Figs. 3–1 and 3–2). Furthermore, the image processing takes place at great speed, often in real time (approximately $\frac{1}{60}$ of a second). Several microcomputer-based image processing sys-

TABLE 3–3 Image Processing Functions Found in Many Image Processing Systems

Preprocessing
A. Radiometric correction (for system and environmental effects)
B. Geometric correction (image to map or image to image)
Display and enhancement
C. Black-and-white display
D. Color composite display
E. Density slice
F. Magnification or reduction
G. Transects
H. Contrast stretch
I. Image algebra (band ratioing, differencing, etc.)
J. Spatial filtering
K. Edge enhancement
L. Principal components
M. Linear combinations (e.g., Kauth transform)
N. Texture transforms
O. Fourier transforms
Information extraction
P. Supervised classification
Q. Unsupervised classification
R. Contextual classification
S. Incorporation of ancillary data in the classification
Geographic information systems (GIS).
T. Raster- or image-based GIS
U. Vector- or polygon-based GIS (must allow polygon comparison)
Integrated systems
W. Complete image processing systems (functions A to S)
Y. Complete image processing systems and GIS (functions A to S and T or U)
Utilities
Z. Mainframe communication for micro- or minicomputer-based systems

Source: Ince (1982) and Jensen (1984).

OFEK06738

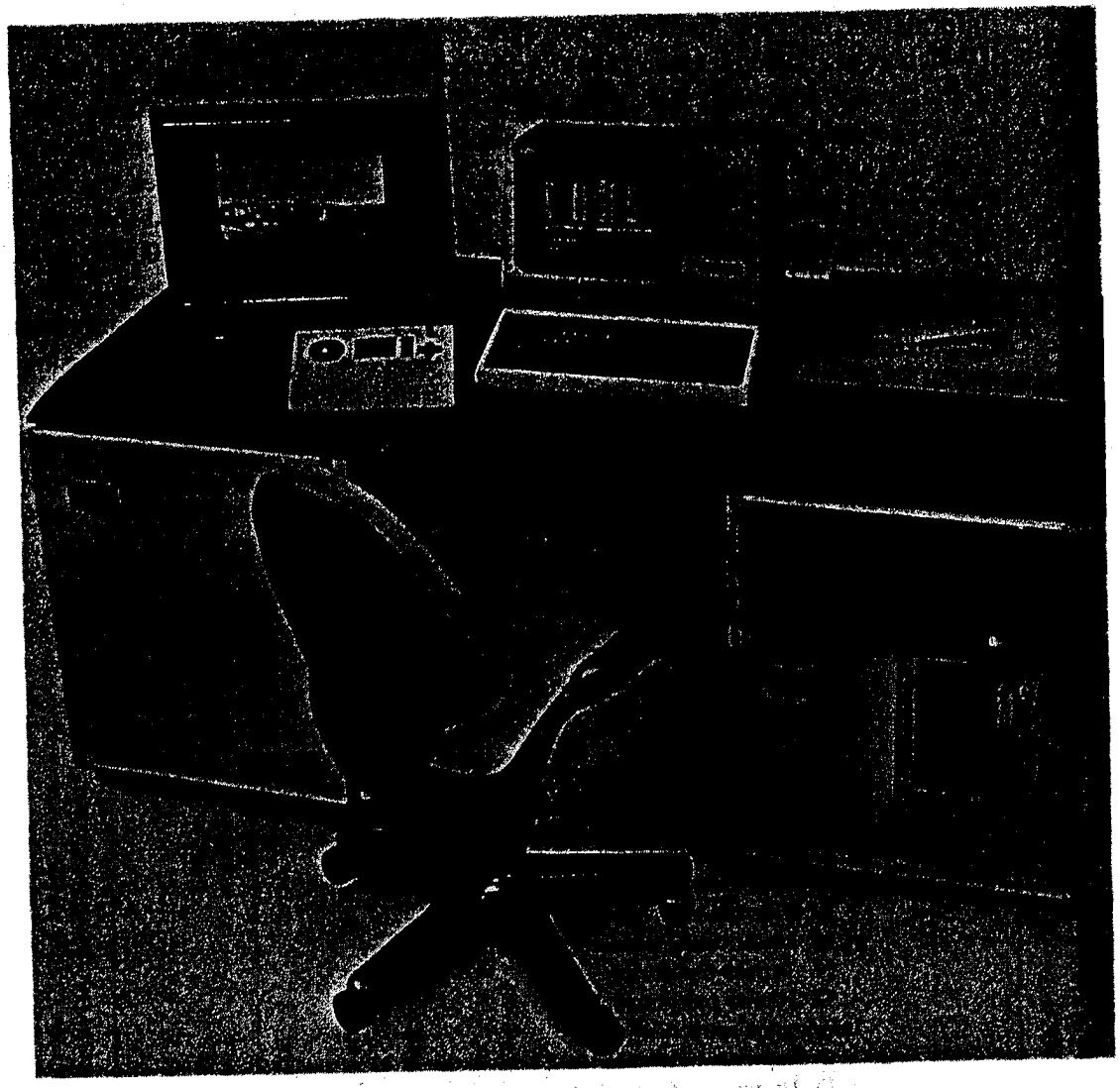


Figure 3–1 Minicomputer-based digital image processing system. A single analyst interacts with the remote sensor data displayed on the high-resolution color monitor using either the trackball or keyboard. The cabinet houses the central processing unit (CPU), the image processor memory used to display the image on the screen, and a 1600-bpi tape drive used to read raw remote sensor data from a computer-compatible tape (CCT). The digitizing tablet is another input device that can be used by the analyst to interact with the image on the CRT. (Courtesy of International Imaging Systems—I²S, Inc.)

tems are now available for under $20,000. These systems provide many of the image processing functions required for both teaching and research (Figs. 3–3 and 3–4). However, the image processing is not performed as rapidly and the display resolution and color resolution of the image processing system (to be discussed) may be much lower.

OFEK06739

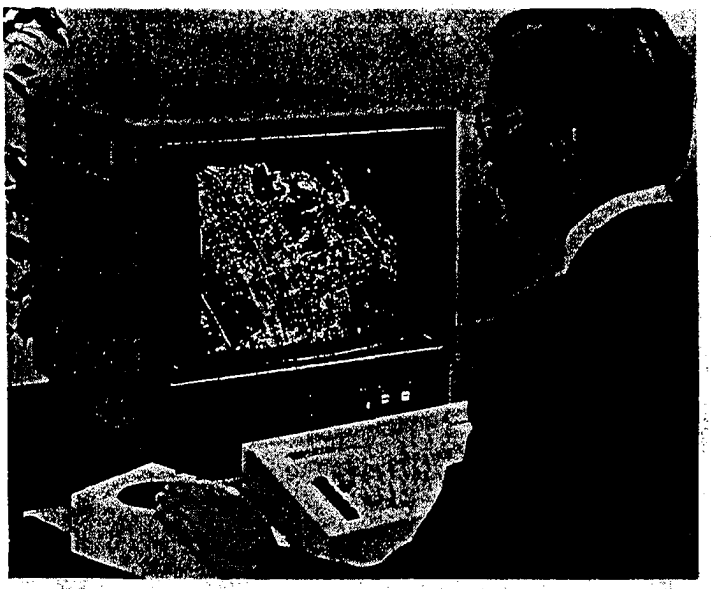


Figure 3–2 Analyst working on a high-resolution digital image processing workstation. The display has 1024 × 1024 display resolution and 24-bit color resolution (i.e., it can display 16.7 million colors at one time out of a palette of 16.7 million colors). The analyst interacts with the image using the trackball or the keyboard. (Courtesy of COMTAL-3M, Inc.)

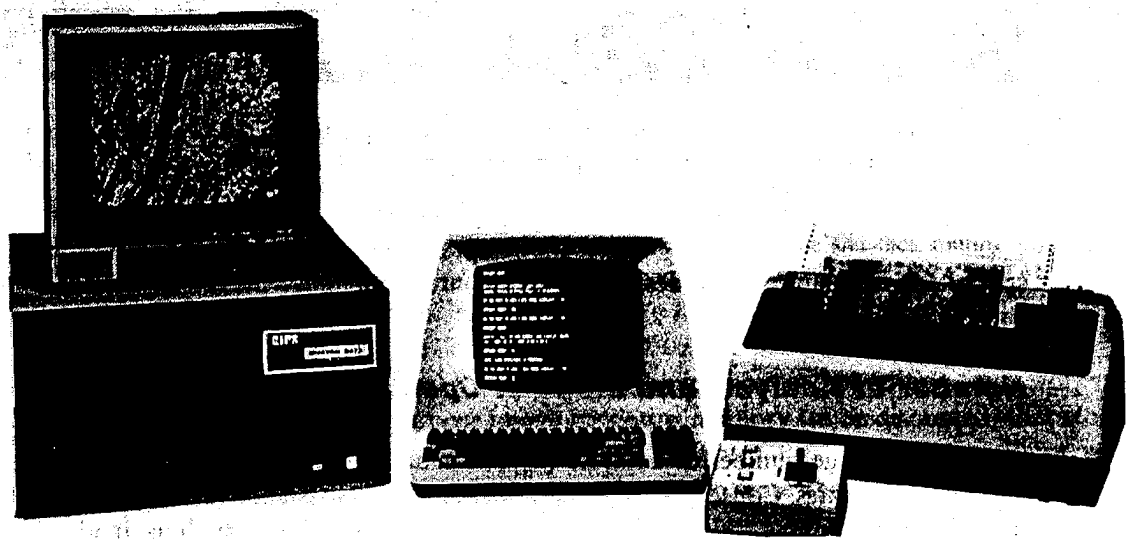


Figure 3–3 Interactive microcomputer-based digital image processing system. This system has a display resolution of 240 × 256 pixels and a color resolution of 12 bits. The maximum number of displayable colors is 4096. (Courtesy of Spectral Data Corporation.)

IMAGE PROCESSING SYSTEM CHARACTERISTICS

When working with or selecting a digital image processing system, the following factors should be considered: the number of analysts to have access to the system at one time, the mode of operation, the central processing unit (CPU), the operating system and type of compiler(s), the amount and type of mass storage required, the

OFEK06740

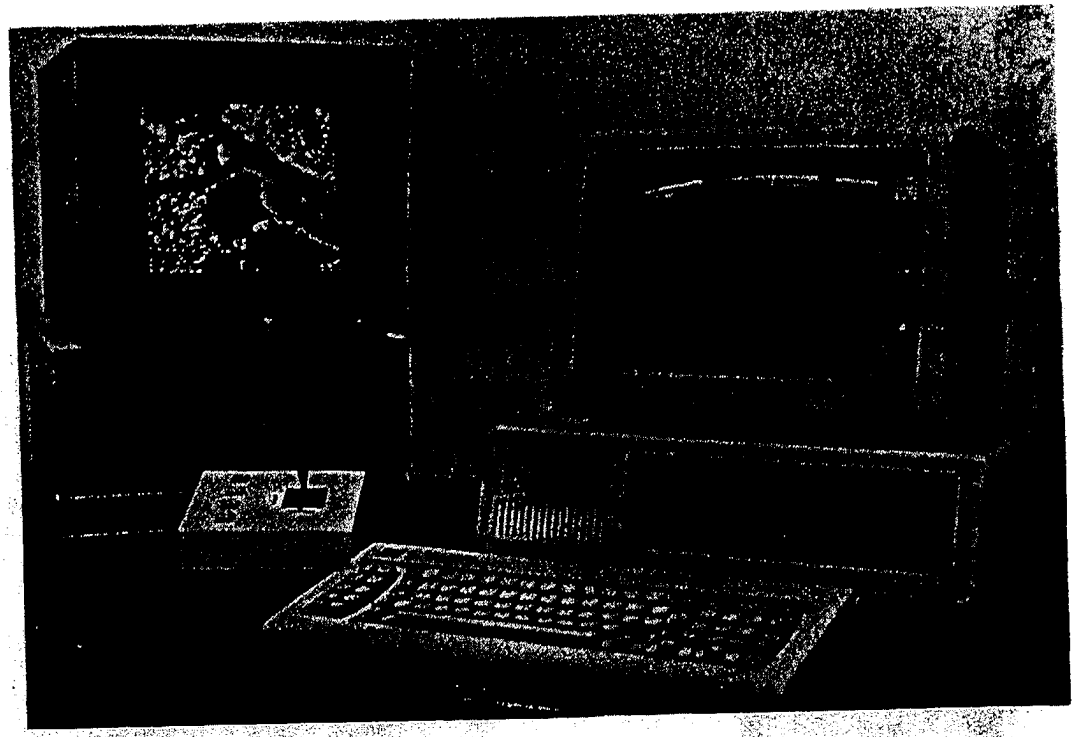


Figure 3–4 Interactive microcomputer-based digital image processing system. This system has a display resolution of 512 × 512 pixels and a color resolution of 24 bits. (Courtesy of ERDAS Inc.)

spatial and color resolution desired, and of course, the image processing applications software.

Number of Analysts and Mode of Operation

A relatively large number of analysts (e.g., 10) must often have access to the image processing facilities. Consequently, the number of analysts assigned to each workstation may range from one, which is exceptional, to perhaps five, which is inadequate. Furthermore, it is ideal if the image processing takes place in an interactive environment where the analyst can see quickly the results of the image processing (Castleman, 1979; McKeown and Denlinger, 1982). Andrews (1977) pointed out that breakthroughs in student "image understanding" are generally accomplished "by placing a human as intimately in the signal processing loop as is possible and allowing his intuitive capabilities to take over." Therefore, it is not wise to educate analysts in a batch mode with very slow turnaround time if interactive learning environments can be provided.

Three interactive digital image processing workstations networked (i.e., they are all connected) to a single minicomputer-based digital image processing system are shown in Fig. 3–5. This relatively expensive configuration allows the workstations

OFEK06741



Figure 3–5 Three interactive digital image processing workstations networked to a single minicomputer. Each workstation has access to its own bank of image processor memory housed in the cabinet. The analysts interact with the image being displayed using the trackball or the keyboard. The image may be modified directly using the local image processor memory or by accessing the minicomputer central processing unit (CPU) in a time-sharing mode and then routing any new information back through the image processor memory for display. (Courtesy of COMTAL-3M, Inc.)

to share (actually time-share) the CPU of the minicomputer. In addition, each workstation has its own special-purpose image processor memory (to be discussed) that stores the remotely sensed data that are displayed on the CRT screen. This allows very rapid digital image processing to take place. Also note that each workstation has access to a trackball. Trackballs or joysticks are used by analysts to interact with the image displayed on the CRT screen.

Central Processing Unit, Arithmetic Coprocessor, and Random Access Memory

The processing of digital remote sensor data requires a large number of central processing unit (CPU) operations. Specifically, the CPU is burdened with two major tasks: numerical calculations and input/output to peripheral mass storage devices, color monitors, and so on. Therefore, it is necessary to have a CPU that can manage data efficiently. In the past, most microcomputer CPUs were designed with 8-bit registers. These 8-bit CPUs were satisfactory for most input/output functions (since

OFEK06742

data are usually transferred 8 bits at a time), but were laboriously slow when numerical calculations were performed. Many microcomputers now have CPUs with 16-bit registers. The 16-bit CPUs perform input/output tasks as fast or faster than did the older 8-bit CPUs. They also compute integer arithmetic expressions at a greater speed than that of their 8-bit predecessors. Even more powerful are the new generation of 32-bit microcomputer CPUs that can address substantially more memory and in many instances operate in a "parallel" fashion (Ince, 1982).

To further enhance the speed of the numerical calculations, an arithmetic (or floating point) coprocessor is often used. The coprocessor works in conjunction with the CPU to perform real number calculations very rapidly. Ideally, an array processor is available that can perform arithmetic operations on an entire matrix of numbers very rapidly. This is especially useful for image enhancement and image analysis operations.

The CPU should contain sufficient random access memory (RAM) for the operating system, image processing applications software, and any remote sensor data that must be held in memory while calculations are performed. Generally, 256K of RAM is sufficient for most image processing when only a few lines of the remote sensor data are held in RAM at one time. However, if the program requires that a large number of lines (e.g., the entire image) be placed in core (RAM) at one time, significantly more RAM is needed.

Operating System and Compiler

The operating system and compiler must be easy to use yet powerful enough so that analysts may program their own relatively sophisticated algorithms and experiment with them on the system. It is not wise to configure an image processing system around an unusual operating system because it becomes difficult to share applications with other scientists. The most universal operating systems are DOS and UNIX.

The compilers most often used in the development of digital image processing software are BASIC, Assembler, and FORTRAN. Very little image processing software has been written in Pascal or PL/1 at the present time (Tables 3–1 and 3–2).

Mass Storage

Digital remote sensor data are usually stored in a matrix format with the various multispectral bands (e.g., blue, green, red, and reflective infrared) in geometric registration one to another (Fig. 2–1). The most common matrix format is band sequential (BSQ), where each spectral band of imagery is stored as an individual file (refer to Chapter 2 for a discussion of this data structure). Each picture element (i.e., pixel) of each band is represented in the computer by a single 8-bit byte (a value from 0 to 255). Over 262,000 bytes of storage (referred to as 262 Kbytes) are required to store a single band of imagery consisting of 512 × 512 pixels. It is often necessary, however, to analyze numerous bands of remote sensing imagery at one time. For example, Landsat thematic mapper imagery contains seven individual bands for each

OFEK06743

geographic area. Therefore, it is not uncommon that more than 1 million bytes of mass storage are required to make the multiple bands of imagery accessible to the CPU for even a small study area. Each pixel of the data matrix must be able to be accessed at random and at great speed.

These two requirements, large storage capacity and random access, are satisfied by the use of disk storage devices. Hard disk systems typically store more than 10 million bytes and provide fast access times. They represent the "ideal" storage medium. In addition to the hard disk, it is desirable to use tape or floppy disk subsystems to provide data input to the system, to back up the hard disk, transfer data between workstations, and archive image data sets or applications software once a project or assignment is completed.

Optical disk storage technology continues to improve and is expected to provide a significant increase in storage capacity at a modest price once the problem of being able to both read from and write to the disk is solved. Optical disks can store gigabytes of data.

Display Resolution

The image processing system should provide low to medium display resolution on the CRT; that is, it should be able to display at least 256 by 256 pixels and preferably more (e.g., 512 by 512 or 1024 by 1024). This allows larger geographic areas to be examined at one time and places the terrain of interest in its regional context. Most geographers and other earth scientists prefer this regional perspective when performing terrain analysis using remote sensor data. Furthermore, it is disconcerting to have to analyze four 256 × 256 images when a single 512 × 512 display provides the information at a glance. Note that some of the microcomputer-based image processing systems listed in Tables 3–1 and 3–2 provide relatively low display resolution.

Color Resolution

This refers to the number of gray-scale shades and/or colors (e.g., 256) which may be displayed on a CRT monitor at one time out of a palette of available colors (e.g., 16.7 million). For many geographic applications, such as high-contrast black-and-white linework cartography, only 1 bit of color is required [i.e., either the line is black or it is white (0 or 1)]. For more sophisticated cartography where many shades of gray or different color combinations are required, up to 8 bits (or 256 colors) may be required. In fact, most thematic mapping applications may be performed quite well by systems that display just 64 user-selectable colors out of a palette of 256 available colors.

The analysis and display of remote sensor image data generally require much higher color resolution than do these cartographic applications. For example, most relatively sophisticated image processing systems (see Table 3–1) provide a tremendous number of displayable colors (e.g., > 4096) out of a large color palette (e.g., 16.7 million). The primary reason for these color requirements is that image analysts

OFEK06744

must often display a "composite" of several individual images at one time on a CRT. This is termed "color compositing." For example, to display a typical color infrared image of Landsat thematic mapper data it is necessary to composite three separate 8-bit images [e.g., the green band (TM 2 = 0.52 to 0.60 μm), the red band (TM 3 = 0.63 to 0.69 μm), and the reflective infrared band (TM 4 = 0.76 to 0.90 μm)]. To obtain a *true-color* or direct-definition image which provides every possible color combination for the three 8-bit images requires that 2^{24} colors (16.7 million) be available in the palette (Fichera and Dickinson, 1983). Such direct-definition systems are ideal but expensive. The reason they are expensive is that every pixel location must be *bit-mapped*. This means that there must be a specific location in memory which keeps track of the exact blue, green, and red color value for every pixel. This requires substantial amounts of computer memory which are usually collected in what is called an *image processor*. Given the availability of image processor memory, the question becomes: What color resolution is adequate?

Generally, 4096 carefully selected colors out of a very large palette (e.g., 16.7 million) appear to be the minimum acceptable for the creation of remote sensing color composites. This provides 12 bits of color with 4 bits available for each of the blue, green, and red image planes (Table 3–4). For image processing applications other than compositing (e.g., black-and-white image display, color density slicing,

TABLE 3–4 Image Processor Memory Required to Produce Various Numbers of Displayable Colors

Image processor memory (bits)	Maximum number of colors displayable at one time on the CRT screen
1	2 (black and white)
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1,024
11	2,048
12	4,096
13	8,192
14	16,384
15	32,768
16	65,536
17	131,072
18	262,144
24	16,777,216

OFEK06745

pattern recognition classification, etc.), the 4096 available colors and large color palette are more than adequate. However, the larger the palette and the greater the number of displayable colors at one time, the better the representation of the remote sensor data on the CRT screen for visual analysis. More will be said in Chapter 5 about how images are displayed using an image processor.

SOFTWARE

Image processing applications software normally resides as executable command modules on the hard disk. The software is usually interactive, prompting the user for specified responses. The programs evaluate the user's response and provide real-time error messages if an inappropriate response is given. This characteristic is essential for some first-time users (Jensen et al., 1979; Jensen, 1983). Programs found in many of the commercially and publicly available image processing systems are summarized in Tables 3–1 and 3–2. This book summarizes the logic and function of many of the more useful image processing functions.

Introduction to the SENSOR Image Processing System

As discussed previously, the ideal is to provide an interactive minicomputer or microcomputer-based image processing system complete with an image processor and high-resolution color monitor. Since this is not always possible, the author has provided the SENSOR image processing system, which will operate in batch mode on most IBM or IBM-compatible mainframes. The software is written in FORTRAN IV, Level G, and the source programs are listed in Appendix A. They may be copied and used freely. These programs provide the reader with a fundamental image processing capability. The graphic output from the SENSOR image processing software is usually to an electrostatic plotter. Instructions for implementing the programs are given in Appendix A.

REFERENCES

ANDREWS, H. C. 1977. "An Educational Digital Image Processing Facility," *Photogrammetric Engineering and Remote Sensing*, Vol. 43, pp. 1161–1168.

CASTLEMAN, K. R. 1979. *Digital Image Processing*. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

EYTON, J. R. 1983. "A Hybrid Image Classification Instruction Package," *Photogrammetric Engineering and Remote Sensing*, Vol. 49, pp. 1175–1181.

FICHERA, R., and D. DICKINSON. 1983. "24-Bit Color Graphics," *Computer Graphics World*, Vol. 6, pp. 25–32.

OFEK06746

INCE, F. 1982. "Digital Image Processing Systems and Remote Sensing," *The International Journal of Remote Sensing*, Vol. 4, pp. 129–148.

JENSEN, J. R. 1983. "Educational Image Processing: An Overview," *Photogrammetric Engineering and Remote Sensing*, Vol. 49, pp. 1151–1157.

JENSEN, J. R. 1984. "Section 9: Recent Developments in the Use of Remote Sensing for Earth Resource Mapping," *The American Cartographer*, Vol. 11, pp. 89–100.

JENSEN, J. R., F. A. ENNERSON, and E. J. HAJIC. 1979. "An Interactive Image Processing System for Remote Sensing Education," *Photogrammetric Engineering and Remote Sensing*, Vol. 45, pp. 1519–1527.

MCKEOWN, D. M., and J. L. DENLINGER. 1982. "Graphical Tools for Interactive Image Interpretation," *Computer Graphics*, Vol. 16, pp. 189–198.

WELCH, R. A., T. R. JORDAN, and E. L. USERY. 1983. "Microcomputers in the Mapping Sciences," *Computer Graphics World*, Vol. 6, pp. 33–42.

WILLIAMS, T. H. L., C. GUNN, and J. SIEBERT. 1983. "Instructional Use of a Mainframe Interactive Image Analysis System," *Photogrammetric Engineering and Remote Sensing*, Vol. 49, pp. 1159–1165.

OFEK06747

4

Initial Statistics Extraction

INTRODUCTION

Once the digital remote sensor data are extracted from the computer-compatible tape (CCT), it is useful to calculate some fundamental univariate and multivariate statistics of the multispectral data. This normally involves computing the minimum and maximum value for each band of imagery, the mean, the standard deviation, a variance-covariance matrix, a correlation matrix, and frequencies of brightness values in each channel which are used to produce histograms. Such statistics provide valuable information necessary for displaying and analyzing remote sensor data (Short, 1982).

UNIVARIATE AND MULTIVARIATE STATISTICS COMPUTED FROM REMOTELY SENSED IMAGERY

Throughout the remainder of this text we will use the following notation:

i = row (or line) in the imagery
j = column (or sample) in the imagery
k = band of imagery
l = another band of imagery
BV_{ijk} = brightness value in a row i, column j, of a band k
BV_r = range of the brightness values (e.g., if $BV_r = 2^8$, then brightness value, BV_{ijk}, may range from 0 to 255)
n = total number of picture elements (pixels) in an array

0950

OFEK06748

μ_k = mean of a band of imagery
s_k = standard deviation of a band of imagery
Var_k = variance of a band of imagery
Cov_{kl} = covariance between pixel values in two bands, k and l
r_{kl} = correlation between pixel values in two bands, k and l
X_c = measurement vector for class c composed of brightness values (BV_{ijk}) from row i, column j, and band k
M_c = mean vector for class c
M_d = mean vector for class d
μ_{ck} = mean value of the data in class c, band k
s_{ck} = standard deviation of the data in class c, band k
V_{ckl} = covariance matrix of class c for bands k through l; often shown as V_c
V_{dkl} = covariance matrix of class d for bands k through l; often shown as V_d

The mean, standard deviation, and variance are useful measures of central tendency [refer to Hammond and McCullagh (1980) if you are not familiar with how they are computed]. However, even more useful from a remote sensing standpoint are measures of covariation and correlation among the several bands. They can provide insight into data redundancy or quality (i.e., are the data of any value for the remote sensing task at hand?). Later it will be shown that variance-covariance and correlation matrices are used in remote sensing principal components analysis (PCA), feature selection, and classification. For this reason, we will closely examine how the variance-covariance between bands is computed and then proceed to compute the correlation between bands. Although performed on a simple data set (consisting of just five pixels of green vegetation from a hypothetical MSS data set), this example should provide insight into the utility of these statistics. Later, the same statistics will be computed for a six-band Charleston, South Carolina, thematic mapper scene consisting of 240 × 256 pixels.

The following examples are based on an analysis of the first five pixels [(1, 1), (1, 2), (1, 3), (1, 4), and (1, 5)] in a four-band (green, red, reflective infrared, reflective infrared) multispectral data set. Thus each pixel consists of four spectral measurements (Table 4–1). Note the low brightness values in band 2 caused by plant chlorophyll absorption of red light for photosynthetic purposes. Increased reflectance of the inci-

TABLE 4–1 Sample Data Set of Brightness Values Used to Demonstrate the Computation of the Variance-Covariance Matrix

Pixel	Band 1 (green)	Band 2 (red)	Band 3 (infrared)	Band 4 (infrared)
(1, 1)	130	57	180	205
(1, 2)	165	35	215	255
(1, 3)	100	25	135	195
(1, 4)	135	50	200	220
(1, 5)	145	65	205	235

OEEK06749

TABLE 4–2 Univariate Statistics for the Sample Data Set

Band	1	2	3	4
Mean μ_k	135	46.4	187	222
Std. dev., s_k	23.71	16.27	31.74	23.87
Variance, Var_k	562.5	264.8	1007.5	570
Minimum	100	25	135	195
Maximum	165	65	215	255

dent infrared energy by the green plant results in higher brightness values in the two infrared bands (3 and 4). Although it is a small sample data set, it represents well the spectral characteristics of healthy green vegetation.

The simple univariate statistics for such data are usually reported in tabular format, as shown in Table 4–2. In this example, band 2 exhibits the smallest variance (264.8), the lowest brightness value (25), the smallest range of brightness values (from 25 to 65), and the lowest mean value (46.4). Conversely, band 3 has the largest variance (1007.5) and the widest range of brightness values (from 135 to 215). These statistics are of value but do not provide any useful information concerning whether or not the spectral measurements in the four bands vary together or are completely independent.

The different remote sensing derived spectral measurements for each pixel often change together in some predictable fashion. If there is no relationship between the brightness value in one band and that in another band for a given pixel, the values are mutually independent; that is, an increase or decrease in one band's brightness value is not accompanied by a predictable change in another band's brightness value. Because spectral measurements of individual pixels may not be independent, some measure of their mutual interaction is needed. This measure, called the *covariance*, is the joint variation of two variables about their common mean. To calculate covariance, we first must compute a quantity similar to a sum of squares. This is called the *corrected sum of products* (SP) and is defined by the equation (Davis, 1973; Davies and Hicks, 1975):

$$SP_{kl} = \sum_{i=1}^{n} (BV_{ik} - \mu_k)(BV_{il} - \mu_l) \tag{4-1}$$

In this notation, BV_{ik} is the ith measurement of band k, and BV_{il} is the ith measurement of band l with n pixels in the study area. The means of bands k and l are μ_k and μ_l, respectively. In our example, variable k might stand for band 1 and variable l could be band 2. It is computationally more efficient to use the following formula to arrive at the same result:

$$SP_{kl} = \sum_{i=1}^{n} (BV_{ik} \times BV_{il}) - \frac{\sum_{i=1}^{n} BV_{ik} \sum_{i=1}^{n} BV_{il}}{n} \tag{4-2}$$

OFEK06750